

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
23 September 2004 (23.09.2004)

PCT

(10) International Publication Number  
**WO 2004/080162 A2**

(51) International Patent Classification<sup>7</sup>:

**A01K**

Leandro; 1311 Victoria Road, Watkinsville, GA 30677 (US).

(21) International Application Number:

PCT/US2004/006378

(74) Agent: **HAYZER, David, J.**; Womble Carlyle Sandridge & Rice, PLLC, P.O. Box 7037, Atlanta, GA 30357-0037 (US).

(22) International Filing Date: 1 March 2004 (01.03.2004)

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

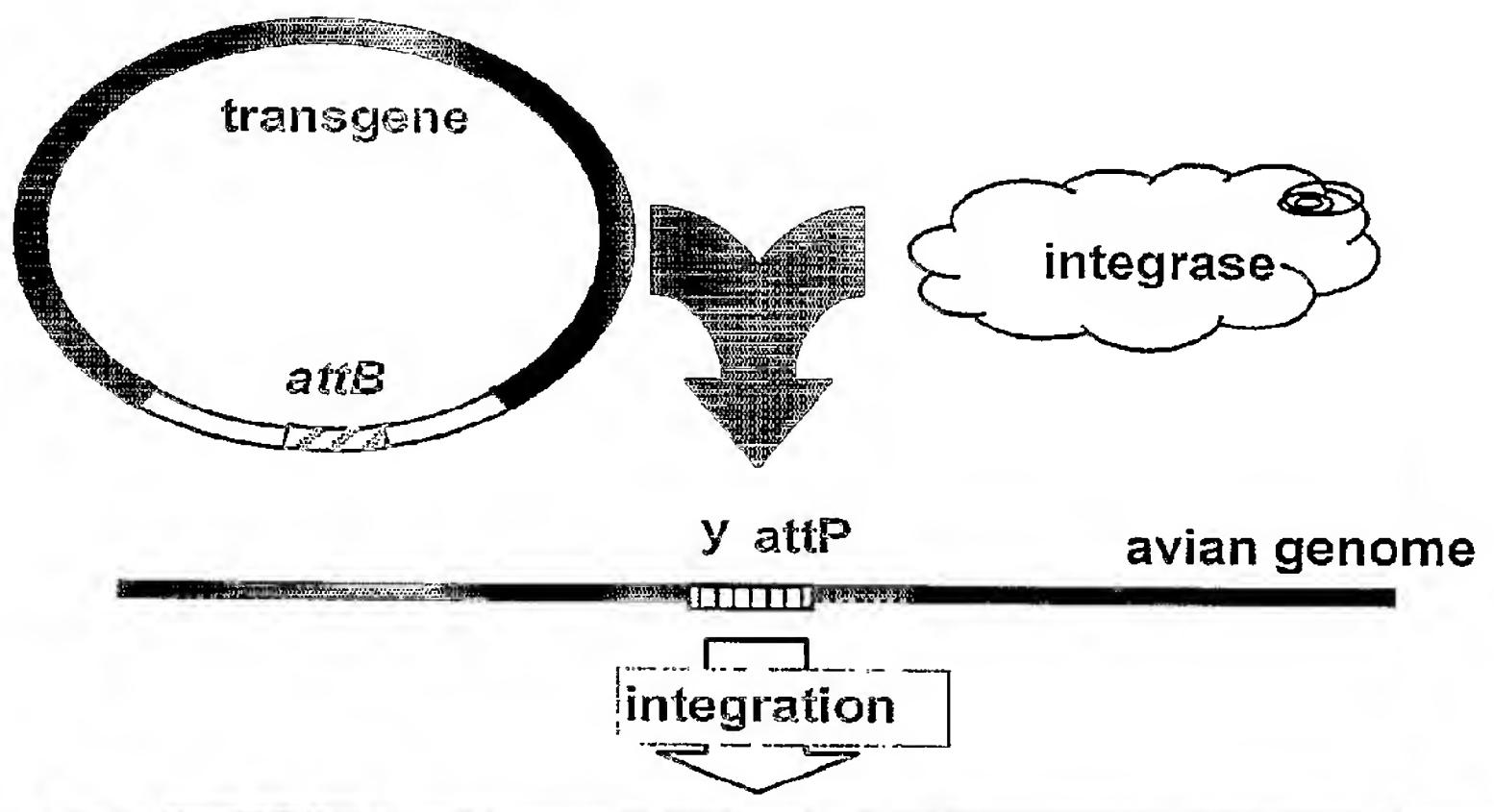
60/453,126	7 March 2003 (07.03.2003)	US
60/490,452	28 July 2003 (28.07.2003)	US
60/536,677	15 January 2004 (15.01.2004)	US

(71) Applicant (for all designated States except US): **AVI-GENICS, INC.** [US/US]; 111 Riverbend Road, Georgia Biobusiness Center, Athens, GA 30605 (US).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK,

[Continued on next page]

(54) Title: INTEGRASE MEDIATED AVIAN TRANSGENESIS



**WO 2004/080162 A2**

(57) **Abstract:** The invention provides methods for integrating a heterologous polynucleotide into the genome of an avian cell. The methods deliver to an avian cell a polynucleotide and a source of integrase activity that mediates recombination between the polynucleotides and the genomic DNA of the avian cell. The invention provides modified avian or artificial chromosomes as vectors to shuttle transgenes or gene clusters into an avian genome. Another aspect of the invention are avian cells genetically modified with a transgene vector. One cell line for the delivery and integration of a transgene comprises a heterologous attP site and, optionally, a region for expressing the integrase. Methods are also included for the production of a heterologous polypeptide by transgenic avian tissue involve integrating a heterologous polynucleotide into the avian genome. Subsequently, a mature transgenic avian is derived by transferring the transgenic blastodermal cells to an embryo and allowing that embryo to develop fully.



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**

- *without international search report and to be republished upon receipt of that report*

**Title of the Invention**

**INTEGRASE MEDIATED AVIAN TRANSGENESIS**

5        The present application claims priority from U.S. provisional patent application Serial Nos. 60/453,126, filed March 7, 2003, 60/490,452, filed July 28, 2003, and 60/536,677 filed January 15, 2004 and which are hereby incorporated by reference in their entireties.

10

**Field of the Invention**

The present invention relates to the field of biotechnology, and more specifically to the field of avian genome modification. Disclosed herein are compositions, vectors, and methods of use thereof, for the generation of genetically transformed avian cells and transgenic birds.

15

**Background**

Transgenic technology to convert animals into "bioreactors" for the production of specific proteins or other substances of pharmaceutical interest (Gordon *et al.*, 1987, *Biotechnology* 5: 1183-1187; Wilmut *et al.*, 1990, *Theriogenology* 33: 113-123) 20 offers significant advantages over more conventional methods of protein production by gene expression. Recombinant nucleic acid molecules, for instance, have been engineered and incorporated into transgenic animals so that an expressed heterologous protein may be joined to a protein or peptide that allows secretion of the transgenic expression product into milk or urine, from which the protein may then be recovered. 25 These procedures, however, may require lactating animals, with the attendant costs of maintaining individual animals or herds of large species, such as cows, sheep, or goats.

Historically, transgenic animals have been produced almost exclusively by microinjection of the fertilized egg. The pronuclei of fertilized eggs are microinjected 30 *in vitro* with foreign, i.e., xenogeneic or allogeneic, heterologous DNA or hybrid

DNA molecules. The microinjected fertilized eggs are then transferred to the genital tract of a pseudopregnant female (e.g., Krimpenfort *et al.*, U.S. Pat. No. 5,175,384).

One system that holds potential is the avian reproductive system. The production of an avian egg begins with formation of a large yolk in the ovary of the hen. The unfertilized oocyte or ovum is positioned on top of the yolk sac. After ovulation, the ovum passes into the infundibulum of the oviduct where it is fertilized if sperm are present, and then moves into the magnum of the oviduct, which is lined with tubular gland cells. These cells secrete the egg-white proteins, including ovalbumin, lysozyme, ovomucoid, conalbumin and ovomucin, into the lumen of the magnum where they are deposited onto the avian embryo and yolk. The hen oviduct offers outstanding potential as a protein bioreactor because of the high levels of protein production, the promise of proper folding and post-translation modification of the target protein, the ease of product recovery, and the shorter developmental period of chickens compared to other potential animal species.

One method for creating permanent genomic modification of an eukaryotic cell is to integrate an introduced DNA into an existing chromosome. Only retroviruses have so far provided efficient integration. However, retroviral integration is directed to a number, albeit limited, of insertion sites within the recipient genome so that positional variation in heterologous gene expression can be evident. Unpredictability as to which insertion site is targeted introduces an undesirable lack of control over the procedure. An additional limitation of the use of retroviruses is that the size of the nucleic acid molecule encoding the virus and heterologous sequences is restricted to about 8 kb. Although wild-type adeno-associated virus (AAV) often integrates at a specific region in the human genome, vectors derived from AAV do not integrate site-specifically due to the deletion of the toxic *rep* gene. Other well-known methods for genomic modification of animal cells include transfection of DNA using calcium phosphate co-precipitation, electroporation, lipofection, microinjection, protoplast fusion and particle bombardment, all of which methods typically produce random integration and at low frequency. Homologous recombination produces site-specific integration, but the frequency of such integration usually is very low.

An alternative method that has been considered for driving the integration of heterologous nucleic acid fragments into a chromosome is the use of a site-specific recombinase (integrase) that can catalyze the insertion or excision of nucleic acid fragments. These enzymes recognize relatively short unique nucleic acid sequences 5 that serve for both recognition and recombination. Examples include Cre (Sternberg & Hamilton, 1981, *J. Mol. Biol.* 150: 467-486, 1981), Flp (Broach *et al.*, 1982, *Cell* 29: 227-234, 1982) and R (Matsuzaki *et al.*, 1990, *J. Bact.* 172: 610-618, 1990).

A novel class of phage integrases that includes the integrase from the phage phiC31 can mediate highly efficient integration of transgenes in mammalian cells both 10 *in vitro* and *in vivo* (Thyagarajan *et al.*, *Mol. Cell Biol.* 21: 3926-3934 (2001)). Constructs and methods of using recombinase to integrate heterologous DNA into a plant, insect or mammalian genome are described by *Calos* in U.S. Patent Serial No. 6,632,672.

The phiC31 integrase is a member of a subclass of integrases, termed serine 15 recombinases, that include R4 and TP901-1. Unlike the phage lambda integrases, which belong to a tyrosine class of recombinases, the serine integrases do not require cofactors such as integration host factor. The phiC31 integrase normally mediates integration of the phiC31 bacteriophage into the genome of *Streptomyces* via recombination between the attP recognition sequence of the phage genome and the 20 attB recognition sequence within the bacterial genome. When a plasmid is equipped with a single attB site, phiC31 integrase will detect and mediate crossover between the attB site and a pseudo-attP site within the mammalian genome. Such pseudo-attP integration sites have now been identified in the mouse and human genomes. If the heterologous DNA is in a circular or supercoiled form, the entire plasmid becomes 25 integrated with *attL* and *attR* arms flanking the nucleic acid insert. PhiC31 integrase is not able to mediate the integration into genomic DNA of sequences bearing attP sites.

PhiC31 integrase-mediated integration results in the destruction of the 30 recognition or recombination sites themselves so that the integration reaction is irreversible. This will bypass the primary concern inherent with other recombinases, i.e., the reversibility of the integration reaction and excision of the inserted DNA.

It has been estimated that there are 50 to 100 pseudo-attP sites in mammalian genomes (mouse and human) and some sites are apparently preferred for integration over others. The chicken genome, however, is only about one-third the size of mammalian genomes, and it was unknown whether there would be a sufficient 5 number of pseudo attP sites in the chicken genome to allow efficient integrase-mediated integration.

We have found that the phiC31 integrase is active in avian cells, increasing the rate of integration over that of a non-integrase-mediated integration. Furthermore, we have determined that the phiC31 integrase works well at both 37° Celsius and 41° 10 Celsius, showing that it will function in the environment of a developing avian embryo.

A need still exists, however, for methods by which avian chromosomes can be permanently modified in an efficient and site-specific manner and the genetically transformed cells used to generate transgenic birds.

15

#### Summary of the Invention

Integration of a transgene into a defined chromosomal site is useful to improve the predictability of expression of the transgene, which is particularly advantageous when creating transgenic avians. Transgenesis by methods that randomly insert a 20 transgene into an avian genome is often inefficient since the transgene may not be expressed at the desired levels or in desired tissues.

A novel class of phage integrases, and in particular the integrase from phage phiC31, can mediate the efficient integration of transgenes into target cells both *in vitro* and *in vivo*. When a plasmid is equipped with a single attB site, phiC31 25 integrase detects attP homologous sequences, termed pseudo-attP sites, in a target genome and mediates crossover between the attB site and a pseudo attP site.

The present invention provides novel methods and recombinant polynucleotide molecules for transfecting and integrating a heterologous nucleic acid molecule into the genome of an avian cell. The methods of the invention deliver to an avian cell 30 population a first nucleic acid molecule that comprises a region encoding a bacterial

recombination site. A source of integrase activity also delivered top the avian cell can be an integrase-encoding nucleic acid sequence and its associated promoter included in the first nucleic acid molecule or as a region of a second nucleic acid molecule that may be co-delivered with the polynucleotide molecule. Alternatively, integrase 5 protein itself can be delivered directly to the target cell.

The recombinant nucleic acid molecules of the present invention may further comprise a heterologous nucleotide sequence operably linked to a promoter so that the heterologous nucleotide sequence, when integrated into the genome DNA of a recipient avian cell, can be expressed to yield a desired polypeptide. The nucleic acid 10 molecule may also include a second transcription initiation site, such as an internal ribosome entry site (IRES), operably linked to a second heterologous polypeptide-encoding region desired to be expressed with the first polypeptide in the same cell.

The heterologous nucleic acid molecule of the present invention may include a cassette for the expression in a recipient avian cell of a desired heterologous 15 polypeptide. Optionally, the nucleic acid molecules may further comprise a marker such as, but not limited to, a puromycin resistance gene, a luciferase gene, EGFP-encoding gene, and the like.

Once delivered to a recipient avian cell, the phiC31 integrase mediates recombination between the att site within the nucleic acid molecule and a 20 bacteriophage attachment site within the genomic DNA of the avian cell. Both att sites are disrupted and the nucleic acid molecule, with partial att sequences at each end, is stably integrated into the genome attP site. The phiC31 integrase, by disrupting the att sites of the incoming nucleic acid and of the recipient site within the avian cell genome, precludes any subsequent reverse recombination event that would 25 excise the integrated nucleic acid and reduce the overall efficiency of stable incorporation of the heterologous nucleic acid.

Following delivery of the nucleic acid molecule and a source of integrase activity into an avian cell population and integrase-mediated recombination, the cells may be returned to an embryo. Late stage blastodermal cells may be returned to a 30 hard shell egg, which is resealed for incubation until hatching. Stage I cells may be directly microinjected with the polynucleotide and source of integrase activity, or

isolated, transfected and returned to a stage I embryo which is reimplanted into a hen for further development. Alternatively, the transfected cells may be maintained in *in vitro* culture.

The present invention further provides modified isolated avian or artificial chromosomes useful as vectors to shuttle transgenes or gene clusters into the avian genome. By delivery to the modified chromosome to an isolated recipient cell, the target cell, and progeny thereof, become trisomic. The additional or trisomic chromosome will not affect the subsequent development of the recipient cell and/or an embryo, nor interfere with the reproductive capacity of an adult bird developed from such cells or embryos. The chromosome will also be stable within chicken cells. The invention provides methods to isolate a population of chromosomes for delivery into chicken embryos or early cells.

The method comprises inserting a lac-operator sequence into an isolated chromosome and, optionally, inserting a desired transgene sequence within the same chromosome. The lac operator region is typically a concatamer of a plurality of lac operators for the binding of multiple lac repressor molecules. A recombinant DNA molecule is constructed that includes an identified region of the target chromosome, a recombination site such as attB or attP, and the lac-operator concatamer. The recombinant molecule is delivered to an avian cell, and homologous recombination will integrate the heterologous polynucleotide and the lac-operator concatamer into the targeted chromosome. A tag-polypeptide, such as the GFP-lac-repressor fusion protein, binds to the lac-operator sequence for identification and isolation of the genetically modified chromosome. The tagged mitotic chromosome can be isolated using, for instance, flow cytometry.

Another aspect of the present invention is an avian cell genetically modified with a transgene vector by the methods of the invention. For example, in one embodiment, the transformed cell can be a chicken early stage blastodermal cell or a genetically transformed cell line, including a sustainable cell line. The transfected cell may comprise a transgene stably integrated into the nuclear genome of the recipient cell, thereby replicating with the cell so that each progeny cell receives a copy of the transfected nucleic acid. A particularly useful cell line for the delivery and integration

of a transgene comprises a heterologous attP site that can increase the efficiency of integration of a polynucleotide by phiC31 integrase and, optionally, a region for expressing the integrase.

Another aspect of the present invention is methods of expressing a heterologous polypeptide in an avian cell by stably transfecting a cell by using site-specific integrase-mediation and a recombinant nucleic acid molecule, as described above, and culturing the transfected cell under conditions suitable for expression of the heterologous polypeptide under the control of the avian transcriptional regulatory region.

Yet another aspect of the present invention concerns transgenic birds, such as chickens, comprising a recombinant nucleic acid molecule and which preferably (though optionally) express a heterologous gene in one or more cells in the animal. Embodiments of the methods for the production of a heterologous polypeptide by the avian tissue involve providing a suitable vector and introducing the vector into embryonic blastodermal cells together with an integrase, preferably phiC31 integrase, so that the vector can integrate into the avian genome. A subsequent step involves deriving a mature transgenic avian from the transgenic blastodermal cells by transferring the transgenic blastodermal cells to an embryo and allowing that embryo to develop fully, so that the cells become incorporated into the bird as the embryo is allowed to develop. An alternative is to transfer a transfected nucleus to an enucleated recipient cell which may then develop into a zygote and ultimately an adult bird. The resulting chick is then grown to maturity.

In various embodiments of the transgenic bird of the present invention, the expression of the transgene may be restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, *trans*-acting factors acting on the transcriptional regulatory region operably linked to the polypeptide-encoding region of interest of the present invention and which control gene expression in the desired pattern. Tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example,

conditional recombination systems or prokaryotic transcriptional regulatory sequences.

The invention can be used to express, in large yields and at low cost, a wide range of desired proteins including those used as human and animal pharmaceuticals, 5 diagnostics, and livestock feed additives. Proteins such as growth hormones, cytokines, structural proteins and enzymes including human growth hormone, interferon, lysozyme, and  $\beta$ -casein are examples of proteins which are desirably expressed in the oviduct and deposited in eggs according to the invention.

Additional objects and aspects of the present invention will become more 10 apparent upon review of the detailed description set forth below when taken in conjunction with the accompanying figures, which are briefly described as follows.

#### Brief Description of the Figures

Fig. 1 illustrates phage integrase-mediated integration. A plasmid vector 15 bearing the transgene includes the attB recognition sequence for the phage integrase. The vector along with integrase-coding mRNA, a vector expressing the integrase, or the integrase protein itself, are delivered into cells or embryos. The integrase recognizes DNA sequences in the avian genome similar to attP sites, termed pseudo-attP, and mediates recombination between the attB and pseudo-attP sites, resulting in 20 the permanent integration of the transgene into the avian genome.

Fig. 2 illustrates the persistent expression of luciferase from a nucleic acid molecule after phiC31 integrase-mediated integration into chicken cells.

Fig. 3 illustrates the results of a puromycin resistance assay to measure phiC31 integrase-mediated integration into chicken cells.

25 Fig. 4 illustrates phiC31 integrase-mediated integration into quail cells. Puromycin resistance vectors bearing attB sites were cotransfected with phiC31 integrase, or a control vector, into QT6 cells, a quail fibrosarcoma cell line. One day after transfection, puromycin was added. Puromycin resistant colonies were counted 12 days post-transfection.

30 Figs. 5A and 5B illustrate that phiC31 integrase can facilitate multiple integrations per avian cell. A puromycin resistance vector bearing an attB site was

cotransfected with an enhanced green fluorescent protein (EGFP) expression vector bearing an attB site, and a phiC31 integrase expression vector. After puromycin selection, many puromycin resistant colonies expressed EGFP in all of their cells. Figs. 5A and 5B are the same field of view with EGFP illuminated with ultraviolet 5 light (Fig. 5A) and puromycin resistant colonies photographed in visible light (Fig. 5B). In Fig. 5B, there are 4 puromycin resistant colonies, two of which are juxtaposed at the top. One of these colonies expressed EGFP.

Fig. 6 shows maps of the small vectors used for integrase assays.

Fig. 7 shows integrase promotes efficient integration of large transgenes in 10 avian cells.

Fig. 8 shows maps of large vectors used for integrase assays.

Fig. 9 illustrates the nucleotide sequence of the integrase-expressing plasmid pCMV-31int (SEQ ID NO: 1).

Fig. 10 illustrates the nucleotide sequence of the plasmid pCMV-luc-attB 15 (SEQ ID NO: 2).

Fig. 11 illustrates the nucleotide sequence of the plasmid pCMV-luc-attP (SEQ ID NO: 3).

Fig. 12 illustrates the nucleotide sequence of the plasmid pCMV-pur-attB (SEQ ID NO: 4).

Fig. 13 illustrates the nucleotide sequence of the plasmid pCMV-pur-attP 20 (SEQ ID NO: 5).

Fig. 14 illustrates the nucleotide sequence of the plasmid pCMV-EGFP-attB (SEQ ID NO: 6).

Fig. 15 illustrates the nucleotide sequence of the plasmid p12.0-lys- 25 LSPIPNMM-CMV-pur-attB (SEQ ID NO: 7).

Fig. 16 illustrates the nucleotide sequence of the plasmid pOMIFN-Ins-CMV- pur-attB (SEQ ID NO: 8).

Fig. 17 illustrates the nucleotide sequence of the integrase-expressing plasmid pRSV-Int (SEQ ID NO: 9).

Fig. 18 illustrates the nucleotide sequence of the plasmid pCR-XL-TOPO- 30 CMV-pur-attB (SEQ ID NO: 10).

Fig. 19 illustrates the nucleotide sequence of the attP containing polynucleotide SEQ ID NO: 11.

Fig. 20 illustrates in schematic form the integration of a heterologous att recombination site into an isolated chromosome. The attB sequence is linked to 5 selectable marker such as a puromycin expression cassette and is flanked by sequences found in the target site of the chromosome to be modified. The DNA is transfected into cells containing the chromosome and stable transfectants are selected by drug resistance. Site specific integration may be confirmed by several techniques including PCR.

10 Fig. 21 illustrates the persistent expression of luciferase from a nucleic acid molecule after phiC31 integrase-mediated integration into chicken cells bearing a wild-type attP sequence.

#### Detailed Description of the Preferred Embodiments

15 This description uses gene nomenclature accepted by the Cucurbit Genetics Cooperative as it appears in the *Cucurbit Genetics Cooperative Report* 18:85 (1995), which are incorporated herein by reference in its entirety. Using this gene nomenclature, genes are symbolized by italicized Roman letters. If a mutant gene is recessive to the normal type, then the symbol and name of the mutant gene appear in 20 italicized lower case letters.

The disclosures of publications, patents, and published patent specifications referenced in this application are hereby incorporated by reference into the present disclosure to more fully describe the state of the art to which this invention pertains.

#### Definitions

25 For convenience, definitions of certain terms employed in the specification, examples, and appended claims are collected here.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise. Thus, for example, reference to "an antigen" includes a mixture of two or more such 30 agents.

The term "avian" as used herein refers to any species, subspecies or race of organism of the taxonomic class *aves*, such as, but not limited to chicken, turkey, duck, goose, quail, pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary. The term includes the various known strains of *Gallus gallus*, or 5 chickens, (for example, White Leghorn, Brown Leghorn, Barred-Rock, Sussex, New Hampshire, Rhode Island, Australorp, Minorca, Amrox, California Gray), as well as strains of turkeys, pheasants, quails, duck, ostriches and other poultry commonly bred in commercial quantities. It also includes an individual avian organism in all stages of development, including embryonic and fetal stages. The term "avian" also may 10 denote "pertaining to a bird", such as "an avian (bird) cell."

The term "nucleic acid" as used herein refers to any natural or synthetic linear and sequential array of nucleotides and nucleosides, for example cDNA, genomic DNA, mRNA, tRNA, oligonucleotides, oligonucleosides and derivatives thereof. For ease of discussion, such nucleic acids may be collectively referred to herein as 15 "constructs," "plasmids," or "vectors." The term "nucleic acid" further includes modified or derivatized nucleotides and nucleosides such as, but not limited to, halogenated nucleotides such as, but not only, 5-bromouracil, and derivatised nucleotides such as biotin-labeled nucleotides.

The terms "polynucleotide," "oligonucleotide," and "nucleic acid sequence" 20 are used interchangeably herein and include, but are not limited to, coding sequences (polynucleotide(s) or nucleic acid sequence(s) which are transcribed and translated into polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory or control sequences); control sequences (e.g., translational start and stop codons, promoter sequences, ribosome binding sites, polyadenylation signals, 25 transcription factor binding sites, transcription termination sequences, upstream and downstream regulatory domains, enhancers, silencers, and the like); and regulatory sequences (DNA sequences to which a transcription factor(s) binds and alters the activity of a gene's promoter either positively (induction) or negatively (repression)). No limitation as to length or to synthetic origin are suggested by the terms described 30 above.

As used herein the terms "peptide," "polypeptide" and "protein" refer to a

polymer of amino acids in a serial array, linked through peptide bonds. A “peptide” typically is a polymer of at least two to about 30 amino acids linked in a serial array by peptide bonds. The term “polypeptide” includes proteins, protein fragments, protein analogues, oligopeptides and the like. The term “polypeptides” contemplates 5 polypeptides as defined above that are encoded by nucleic acids, produced through recombinant technology (isolated from an appropriate source such as a bird), or synthesized. The term “polypeptides” further contemplates polypeptides as defined above that include chemically modified amino acids or amino acids covalently or noncovalently linked to labeling moieties.

10 The terms “percent sequence identity” or “percent sequence similarity” as used herein refer to the degree of sequence identity between two nucleic acid sequences or two amino acid sequences as determined using the algorithm of Karlin & Atschul, *Proc. Natl. Acad. Sci.* 87: 2264-2268 (1990), modified as in Karlin & Atschul, *Proc. Natl. Acad. Sci.* 90: 5873-5877 (1993). Such an algorithm is incorporated into the 15 NBLAST and XBLAST programs of Atschul *et al.*, 1990, *T. Mol. Biol.* Q15: 403-410. BLAST nucleotide searches are performed with the NBLAST program, score = 100, word length = 12, to obtain nucleotide sequences homologous to a nucleic acid molecule of the invention. BLAST protein searches are performed with the XBLAST program, score = 50, word length = 3, to obtain amino acid sequences homologous to 20 a reference polypeptide. To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized as described in Atschul *et al.*, *Nucl. Acids Res.* 25: 3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g. XBLAST and NBLAST) are used. Other algorithms, programs and default settings may also be suitable such as, but not only, 25 the GCG-Sequence Analysis Package of the U.K. Human Genome Mapping Project Resource Centre that includes programs for nucleotide or amino acid sequence comparisons. Examples of preferred algorithms are FASTA and BESTFIT.

The terms “recombinant nucleic acid” and “recombinant DNA” as used herein refer to combinations of at least two nucleic acid sequences that are not naturally 30 found in a eukaryotic or prokaryotic cell. The nucleic acid sequences may include, but are not limited to, nucleic acid vectors, gene expression regulatory elements, origins

of replication, suitable gene sequences that when expressed confer antibiotic resistance, protein-encoding sequences and the like. The term "recombinant polypeptide" is meant to include a polypeptide produced by recombinant DNA techniques. A recombinant polypeptide may be distinct from a naturally occurring 5 polypeptide either in its location, purity or structure. Generally, a recombinant polypeptide will be present in a cell in an amount different from that normally observed in nature.

The term "gene" or "genes" as used herein refers to nucleic acid sequences that encode genetic information for the synthesis of a whole RNA, a whole protein, or any 10 portion of such whole RNA or whole protein. Genes that are not naturally part of a particular organism's genome are referred to as "foreign genes," "heterologous genes" or "exogenous genes" and genes that are naturally a part of a particular organism's genome are referred to as "endogenous genes". The term "gene product" refers to an RNA or protein that is encoded by the gene. "Endogenous gene products" are RNAs 15 or proteins encoded by endogenous genes. "Heterologous gene products" are RNAs or proteins encoded by "foreign, heterologous or exogenous genes" and are, therefore, not naturally expressed in the cell.

The term "expressed" or "expression" as used herein refers to the transcription from a gene to give an RNA nucleic acid molecule at least complementary in part to a 20 region of one of the two nucleic acid strands of the gene. The term "expressed" or "expression" as used herein may also refer to the translation from an RNA molecule to give a protein, a polypeptide or a portion thereof.

The term "operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Control 25 sequences operably linked to a coding sequence are capable of effecting the expression of the coding sequence. The control sequences need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. For example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be 30 considered "operably linked" to the coding sequence.

The term "transcription regulatory sequences" as used herein refers to nucleotide sequences that are associated with a gene nucleic acid sequence and which regulate the transcriptional expression of the gene. Exemplary transcription regulatory sequences include enhancer elements, hormone response elements, steroid response elements, negative regulatory elements, and the like.

5 The term "promoter" as used herein refers to the DNA sequence that determines the site of transcription initiation by an RNA polymerase. A "promoter-proximal element" is a regulatory sequence generally within about 200 base pairs of the transcription start site.

10 The term "internal ribosome entry sites (IRES)" as used herein refers to a region of a nucleic acid, most typically an RNA molecule, wherein eukaryotic initiation of protein synthesis occurs far downstream of the 5' end of the RNA molecule. A 43S pre-initiation complex comprising the elf2 protein bound to GTP and Met-tRNA<sub>i</sub><sup>Met</sup>, the 40S ribosomal subunit, and factors elf3 and 3lf1A may bind to  
15 an "IRES" before locating an AUG start codon. An "IRES" may be used to initiate translation of a second coding region downstream of a first coding region, wherein each coding region is expressed individually, but under the initial control of a single upstream promoter. An "IRES" may be located in a eukaryotic cellular mRNA.

20 The term "coding region" as used herein refers to a continuous linear arrangement of nucleotides which may be translated into a polypeptide. A full length coding region is translated into a full length protein; that is, a complete protein as would be translated in its natural state absent any post-translational modifications. A full length coding region may also include any leader protein sequence or any other region of the protein that may be excised naturally from the translated protein.

25 The terms "vector" or "nucleic acid vector" as used herein refer to a natural or synthetic single or double stranded plasmid or viral nucleic acid molecule (RNA or DNA) that can be transfected or transformed into cells and replicate independently of, or within, the host cell genome. The term "expression vector" as used herein refers to a nucleic acid vector that comprises a transcription regulatory region operably linked  
30 to a site wherein is, or can be, inserted, a nucleotide sequence to be transcribed and,

optionally, to be expressed, for instance, but not limited to, a sequence coding at least one polypeptide.

The term "transfection" as used herein refers to the process of inserting a nucleic acid into a host cell. Many techniques are well known to those skilled in the art to facilitate transfection of a nucleic acid into an eukaryotic cell. These methods include, for instance, treating the cells with high concentrations of salt such as a calcium or magnesium salt, an electric field, detergent, or liposome mediated transfection, to render the host cell competent for the uptake of the nucleic acid molecules, and by such methods as micro-injection into a pro-nucleus, sperm-mediated and restriction-mediated integration.

The terms "recombinant cell" and "genetically transformed cell" refer to a cell comprising a combination of nucleic acid segments not found in a single cell with each other in nature. A new combination of nucleic acid segments can be introduced into an organism using a wide array of nucleic acid manipulation techniques available to those skilled in the art. The recombinant cell may harbor a vector that is extragenomic, i.e. that does not covalently insert into the cellular genome, including a non-nuclear (e.g. mitochondrial) genome(s). A recombinant cell may further harbor a vector or a portion thereof that is intragenomic, i.e. covalently incorporated within the genome of the recombinant cell.

As used herein, a "transgenic avian" is any avian, as defined above, including the chicken and quail, in which one or more of the cells of the avian contain heterologous nucleic acid introduced by manipulation, such as by transgenic techniques. The nucleic acid may be introduced into a cell, directly or indirectly, by introduction into a precursor of the cell by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. Genetic manipulation also includes classical cross-breeding, or *in vitro* fertilization. A recombinant DNA molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA.

The terms "chimeric animal" or "mosaic animal" are used herein to refer to animals in which the recombinant gene is found, or in which the recombinant is expressed, in some but not all cells of the animal. The term "tissue-specific chimeric

animal" indicates that the recombinant gene is present and/or expressed in some tissues but not others.

As used herein, the term "transgene" means a nucleic acid sequence that is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout).

10 The term "cytokine" as used herein refers to any secreted polypeptide that affects a function of cells and modulates an interaction between cells in the immune, inflammatory or hematopoietic response. A cytokine includes, but is not limited to, monokines and lymphokines. Examples of cytokines include, but are not limited to, interferon  $\alpha$ 2b, Interleukin-1 (IL-1), Interleukin-6 (IL-6), Interleukin-8 (IL-8), Tumor 15 Necrosis Factor- $\alpha$  (TNF- $\alpha$ ) and Tumor Necrosis Factor  $\beta$  (TNF- $\beta$ ).

20 The term "antibody" as used herein refers to polyclonal and monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof. Antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments,  $F(ab')_2$  fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

25 The term "immunoglobulin polypeptide" as used herein refers to a constituent polypeptide of an antibody or a polypeptide derived therefrom. An "immunological polypeptide" may be, but is not limited to, an immunological heavy or light chain and may include a variable region, a diversity region, joining region and a constant region or any combination, variant or truncated form thereof. The term "immunological polypeptides" further includes single-chain antibodies comprised of, but not limited to, an immunoglobulin heavy chain variable region, an immunoglobulin light chain variable region and optionally a peptide linker.

30 The terms "integrase" and "integrase activity" as used herein refer to a nucleic acid recombinase of the serine recombinase family of proteins.

The term "source of integrase activity" as used herein refers to a polypeptide or multimeric protein having serine recombinase (integrase) activity in an avian cell. The term may further refer to a polynucleotide encoding the serine recombinase, such as an mRNA, an expression vector, a gene or isolated gene that may be expressed as 5 the recombinase-specific polypeptide or protein.

The term "recombination site" as used herein refers to a polynucleotide stretch comprising a recombination site normally recognized and used by an integrase. For example,  $\lambda$  phage is a temperate bacteriophage that infects *E. coli*. The phage has one attachment site for recombination (attP) and the *E. coli* bacterial genome has an 10 attachment site for recombination (attB). Both of these sites are recombination sites for  $\lambda$  integrase. Recombination sites recognized by a particular integrase can be derived from a homologous system and associated with heterologous sequences, for example, the attP site can be placed in other systems to act as a substrate for the integrase.

15 The term "pseudo-recombination site" as used herein refers to a site at which an integrase can facilitate recombination even though the site may not have a sequence identical to the sequence of its wild-type recombination site. For example, a phiC31 integrase and vector carrying a phiC31 wild-type recombination site can be placed into an avian cell. The wild-type recombination sequence aligns itself with a sequence 20 in the avian cell genome and the integrase facilitates a recombination event. When the sequence from the genomic site in the avian cell, where the integration of the vector took place, is examined, the sequence at the genomic site typically has some identity to, but may not be identical with, the wild-type bacterial genome recombination site. The recombination site in the avian cell genome is considered to be a pseudo- 25 recombination site (e.g., a pseudo-attP site) at least because the avian cell is heterologous to the normal phiC31 phage/bacterial cell system. The size of the pseudo-recombination site can be determined through the use of a variety of methods including, but not limited to, (i) sequence alignment comparisons, (ii) secondary structural comparisons, (iii) deletion or point mutation analysis to find the functional 30 limits of the pseudo-recombination site, and (iv) combinations of the foregoing.

A nucleic acid fragment of interest may be a trait-producing sequence, by which it is meant a sequence conferring a non-native trait upon the cell in which the protein encoded by the trait-producing sequence is expressed. The term "non-native" when used in the context of a trait-producing sequence means that the trait produced is 5 different than one would find in an unmodified organism which can mean that the organism produces high amounts of a natural substance in comparison to an unmodified organism, or produces a non-natural substance. For example, the genome of a bird could be modified to produce proteins not normally produced in birds such as, for instance, human or mouse antibodies, human cytokines, etc. Other useful traits 10 include disease resistance, meat flavor, animal size, and the like.

A nucleic acid fragment of interest may additionally be a "marker nucleic acid" or expressed as a "marker polypeptide". Marker genes encode proteins that can be easily detected in transformed cells and are, therefore, useful in the study of those cells. Examples of suitable marker genes include  $\beta$ -galactosidase, green or yellow 15 fluorescent proteins, enhanced green fluorescent protein, chloramphenicol acetyl transferase, luciferase, and the like. Such regions may also include those 5' noncoding sequences involved with initiation of transcription and translation, such as the enhancer, TATA box, capping sequence, CAAT sequence, and the like

The term "transformed" as used herein refers to a heritable alteration in a cell 20 resulting from the uptake of a heterologous DNA.

The term "trisomic" as used herein refers to a cell or animal, such as an avian cell or bird that has a  $2n+1$  chromosomal complement, where  $n$  is the haploid number of chromosomes, for the animal species concerned.

Techniques useful for isolating and characterizing the nucleic acids and 25 proteins of the present invention are well known to those of skill in the art and standard molecular biology and biochemical manuals may be consulted to select suitable protocols without undue experimentation. See, for example, Sambrook et al, 1989, "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor, the content of which is herein incorporated by reference in its entirety.

Abbreviations

Abbreviations used in the present specification include the following: aa, amino acid(s); bp, base pair(s); kb, kilobase; att, bacterial recombination attachment site; IU, infectious units.

5 In the standard method of integrase mediated-transgenesis, a serine recombinase integrase mediates recombination between an attB site on a transgene vector and a pseudo attP site on a chromosome. In the method of the invention for integrase-mediated transgenesis, a heterologous wild-type attP site can be integrated into an avian nuclear genome to create a transgenic cell line or bird. A serine recombinase (integrase) and an attB-bearing transgene vector are then introduced into cells harboring the heterologous attP site, or into embryos derived from birds which bear the attP recombination site. The locations of attP and attB may be reversed such that the attB site is inserted into an avian chromosome and the attP sequence resides in an incoming transgene vector. In either case, the att site of the introduced vector 10 would then preferentially recombine with the integrated heterologous att site in the genome of the recipient cell.

15

The methods of the invention are based, in part, on the discovery that there exist in avian genomes a number of specific nucleic acid sequences, termed pseudo-recombination sites, the sequences of which may be distinct from wild-type 20 recombination sites but which can be recognized by a site-specific integrase and used to promote the efficient insertion of heterologous genes or polynucleotides into the targeted avian nuclear genome. The inventors have identified pseudo-recombination sites in avian cells capable of recombining with a recombination site, such as an attB site within a recombinant nucleic acid molecule introduced into the target avian cell. 25 The invention is also based on the prior integration of a heterologous att recombination site, typically isolated from a bacteriophage or a modification thereof, into the genome of the target avian cell.

Integration into a predicted chromosomal site is useful to improve the predictability of expression, which is particularly advantageous when creating 30 transgenic avians. Transgenesis by methods that result in insertion of the transgene

into random positions of the avian genome is unpredictable since the transgene may not express at the expected levels or in the predicted tissues.

The invention as disclosed herein, therefore, provides methods for site-specifically genetically transforming an avian nuclear genome. In general, an avian cell having a first recombination site in the nuclear genome is transformed with a site-specific polynucleotide construct comprising a second recombination sequence and one or more polynucleotides of interest. Into the same cell, integrase activity is introduced that specifically recognizes the first and second recombination sites under conditions such that the polynucleotide sequence of interest is inserted into the nuclear genome via an integrase-mediated recombination event between the first and second recombination sites.

The integrase activity, or a source thereof, can be introduced into the avian cell prior to, or concurrent with, the introduction of the site-specific construct. The integrase can be delivered to a cell as a polypeptide, or by expressing the integrase from a source polynucleotide such as an mRNA or from an expression vector that encodes the integrase, either of which can be delivered to the target avian cell before, during or after delivery of the polynucleotide of interest. Any integrase that has activity in an avian cell may be useful in the present invention, including HK022 (Kolot *et al.*, *Biotechnol. Bioeng.*, 84: 56-60 (2003)). Preferably, the integrase is a serine recombinase as described, for example, by Smith & Thorpe, in *Mol. Microbiol.*, 44: 299-307 (2002). More preferably, the integrase is a bacteriophage integrase such as, but not limited to, TP901-1 (Stoll *et al.*, *J. Bact.*, 184: 3657-3663 (2002); Olivares *et al.*, *Gene*, 278:167-176 (2001). Most preferably, the integrase is from the phage phiC31.

The nucleotide sequence of the junctions between an integrated transgene into the attP (or attB site) would be known. Thus, a PCR assay can be designed by one of skill in the art to detect when the integration event has occurred. The PCR assay for integration into a heterologous wild-type attB or attP site can also be readily incorporated into a quantitative PCR assay using TAQMANTM or related technology so that the efficiency of integration can be measured.

The minimal attB and attP sites able to catalyze recombination mediated by

the phiC31 integrase are 34 and 39 bp, respectively. In cell lines that harbor a heterologous integrated attP site, however, integrase has a preference for the inserted attP over any pseudo-attP sites of similar length, because pseudo-attP sites have very low sequence identity (between 10 to 50% identity) compared to the more efficient 5 wild-type attP sequence. It is within the scope of the methods of the invention, however, for the recombination site within the target avian genome to be a pseudo-att site such as a pseudo-attP site or an attP introduced into an avian genome.

The sites used for recognition and recombination of phage and bacterial DNAs (the native host system) are generally non-identical, although they typically have a 10 common core region of nucleic acids. The bacterial sequence is generally called the attB sequence (bacterial attachment) and the phage sequence is called the attP sequence (phage attachment). Because they are different sequences, recombination will result in a stretch of nucleic acids (called attL or attR for left and right) that is neither an attB sequence or an attP sequence, and likely is functionally unrecognizable 15 as a recombination site to the relevant enzyme, thus removing the possibility that the enzyme will catalyze a second recombination reaction that would reverse the first.

The integrase may recognize a recombination site where sequence of the 5' region of the recombination site can differ from the sequence of the 3' region of the recombination sequence. For example, for the phage phiC31 attP (the phage 20 attachment site), the core region is 5'-TTG-3' the flanking sequences on either side are represented here as attP5' and attP3', the structure of the attP recombination site is, accordingly, attP5'-TTG-attP3'. Correspondingly, for the native bacterial genomic target site (attB) the core region is 5'-TTG-3', and the flanking sequences on either side are represented here as attB5' and attB3', the structure of the attB recombination 25 site is, accordingly, attB5'-TTG-attB3'. After a single-site, phiC31 integrase-mediated recombination event takes place between the phiC31 phage and the bacterial genome, the result is the following recombination product: attB5'-TTG-attP3'{phiC31 vector sequences}attP5'-TTG-attB3'. In the method of invention, the attB site will be within a recombinant nucleic acid molecule that may be delivered to a target avian cell. The 30 corresponding attP (or pseudo-attP) site will be within the avian cell nuclear genome. Consequently, after phiC31 integrase mediated recombination, the recombination

product, the nuclear genome with the integrated heterologous polynucleotide will have the sequence attP5'-TTG-attB3'{heterologous polynucleotide}-attB5'-TTG-attP3'. Typically, after recombination the post-recombination recombination sites are no longer able to act as substrate for the phiC31 integrase. This results in stable 5 integration with little or no integrase mediated excision.

While the preferred recombination site to be included in the recombinant nucleic acid molecules and modified chromosomes of the present invention is the attP site, it is contemplated that any attP-like site may be used if compatible with the attB site. For instance, any pseudo-attP site of the chicken genome may be identified 10 according to the methods of Example 7 below and used as a heterologous att recombination site. Such attP-like sites may have a sequence that is at least 25% identical to SEQ ID NO: 11 as shown in Fig. 19, such as described in Groth *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 97: 5995-6000 (2000) incorporated herein by reference in its entirety. Preferably the selected site will have at least the same degree of 15 efficiency of recombination as the attP site (SEQ ID NO: 11) itself.

In the methods of the present invention, the recipient avian cell population may be an isolated avian cell line such as, for example, DF-1 chicken fibroblasts, chicken DT40 cells or a cell population derived from an early stage embryo such as a chicken stage I or stage X embryo. A particularly useful avian cell population is 20 blastodermal cells isolated from an early stage I embryo or a stage X avian embryo. The methods of the present invention, therefore, include steps for the isolation of blastodermal cells that are then suspended in a cell culture medium or buffer for maintaining the cells in a viable state, and which allows the cell suspension to contact the nucleic acids of the present invention. It is also within the scope of the invention 25 for the nucleic acid construct and the source of integrase activity to be delivered directly to an avian embryo such as a blastodermal layer, or to a tissue layer of an adult bird such as the lining of an oviduct.

When the recipient avian cell population is isolated from an early stage avian embryo, the embryos must first be isolated. For stage I avian embryos from, for 30 example, a chicken, a fertilized ovum is surgically removed from a bird before the deposition of the outer hard shell has occurred. The nucleic acids for integrating a

heterologous nucleic acid into a recipient avian cell genome may then be delivered to isolated embryos by lipofection, microinjection (as described in Example 6 below) or electroporation and the like. After delivery of the nucleic acid, the transfected embryo and its yolk may be deposited into the infundibulum of a recipient hen for the 5 deposition of egg white proteins and a hard shell, and laying of the egg. Stage X avian embryos are obtained from freshly laid fertilized eggs and the blastodermal cells isolated as a suspension of cells in a medium, as described in Example 4 below. Isolated stage X blastodermal cell populations, once transfected, may be injected into recipient stage X embryos and the hard shell eggs resealed according to the methods 10 described in U.S. Patent No. 6,397,777.

In the methods of the invention, once a heterologous nucleic acid is delivered to the recipient avian cell, the integrase activity is expressed. The expressed integrase (or injected integrase polypeptide) then mediates recombination between the att site of the heterologous nucleic acid molecule, and the att (or pseudo att) site within the 15 genomic DNA of the recipient avian cell.

It is within the scope of the present invention for the integrase-encoding sequence and a promoter operably linked thereto to be included in the delivered nucleic acid molecule and that expression of the integrase activity occurs before integration of the heterologous nucleic acid into the avian cell genome. Preferably, 20 the integrase-encoding nucleic acid sequence and associated promoter are in an expression vector that may be co-delivered to the recipient avian cell with the heterologous nucleic acid molecule to be integrated into the recipient genome.

One suitable integrase expressing expression vector for use in the present invention is pCMV-C31int (SEQ ID NO: 1) as shown in Fig. 9, and described in 25 Groth *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 97: 5995-6000 (2000), incorporated herein by reference in its entirety. In pCMV-C31int, expression of the integrase-encoding sequence is driven by the CMV promoter. However, any promoter may be used that will give expression of the integrase in a recipient avian cell, including operably linked avian-specific gene expression control regions of the avian ovalbumin, 30 lysozyme, ovomucin, ovomucoid gene loci, viral gene promoters, inducible promoters, the RSV promoter and the like.

The recombinant nucleic acid molecules of the present invention for delivery of a heterologous polynucleotide to the genome of a recipient avian cell may comprise a nucleotide sequence encoding the attB attachment site of *Streptomyces ambofaciens* as described in Thorpe & Smith, *Proc. Natl. Acad. Sci. U.S.A.* 95: 5505-5510 (1998).

5 The nucleic acid molecule of the present invention further comprises an expression cassette for the expression in a recipient avian cell of a heterologous nucleic acid encoding a desired heterologous polypeptide. Optionally, the nucleic acid molecules may further comprise a marker such as, but not limited to, a puromycin resistance gene, a luciferase gene, EGFP, and the like.

10 It is contemplated that the expression cassette for introducing a desired heterologous polypeptide comprises a promoter operably linked to a nucleic acid encoding the desired polypeptide and, optionally, a polyadenylation signal sequence. Exemplary nucleic acids suitable for use in the present invention are more fully described in the examples below.

15 In the methods of the present invention, following delivery of the nucleic acid molecule and a source of integrase activity into an avian cell population, the cells are maintained under culture conditions suitable for the expression of the integrase and/or for the integrase to mediate recombination between the recombination site of the nucleic acid and recombination site in the genome of the recipient avian cell. When 20 the recipient avian cell is cultured *in vitro*, such cells may be incubated at 37° Celsius if the cells are chicken early stage blastodermal cells. They may then be injected into an embryo within a hard shell, which is resealed for incubation until hatching. Alternatively, the transfected cells may be maintained in *in vitro* culture.

25 Site-Specific Nucleic Acid Constructs and Methods of Delivery to an Avian Cell

The present invention provides methods for the site-specific insertion of a heterologous nucleic acid molecule into the nuclear genome of an avian cell by delivering to a target avian cell that has a recombination site in its nuclear genome, a source of integrase activity, a site-specific construct that has another recombination 30 site and a polynucleotide of interest, and allowing the integrase activity to facilitate a recombination event between the two recombination sites, thereby integrating the

polynucleotide of interest into the avian nuclear genome.

(a) *Expression vector nucleic acid molecules*: A variety of recombinant nucleic acid expression vectors are suitable for use in the practice of the present invention. The site-specific constructs described herein can be constructed utilizing methodologies 5 well known in the art of molecular biology (see, for example, *Ausubel* or *Maniatis*) in view of the teachings of the specification. As described above, the constructs are assembled by inserting into a suitable vector backbone a recombination site such as an attP or an attB site, a polynucleotide of interest operably linked to a gene expression control region of interest and, optionally a sequence encoding a positive selection 10 marker. Polynucleotides of interest can include, but are not limited to, expression cassettes encoding a polypeptide to be expressed in the transformed avian cell or in a transgenic bird derived therefrom. The site-specific constructs are typically circular and may also contain selectable markers, an origin of replication, and other elements.

Any of the vectors of the present invention may also optionally include a 15 sequence encoding a signal peptide that directs secretion of the polypeptide expressed by the vector from the transgenic cells, for instance, from tubular gland cells of the oviduct. This aspect of the invention effectively broadens the spectrum of exogenous proteins that may be deposited in the whites of avian eggs using the methods of the invention. Where an exogenous polypeptide would not otherwise be secreted, the 20 vector bearing the coding sequence can be modified to comprise, for instance, about 60 bp encoding a signal peptide. The DNA sequence encoding the signal peptide is inserted in the vector such that the signal peptide is located at the N-terminus of the polypeptide encoded by the vector.

The expression vectors of the present invention can comprise an avian 25 transcriptional regulatory region for directing expression of either fusion or non-fusion proteins. With fusion vectors, a number of amino acids are usually added to the desired expressed target gene sequence such as, but not limited to, a polypeptide sequence for thioredoxin. A proteolytic cleavage site may further be introduced at a site between the target recombinant protein and the fusion sequence. Additionally, a 30 region of amino acids such as a polymeric histidine region may be introduced to allow binding of the fusion protein to metallic ions such as nickel bonded to a solid support,

for purification of the fusion protein. Once the fusion protein has been purified, the cleavage site allows the target recombinant protein to be separated from the fusion sequence. Enzymes suitable for use in cleaving the proteolytic cleavage site include, but are not limited to, Factor Xa and thrombin. Fusion expression vectors that may be 5 useful in the present invention include pGex (Amrad Corp., Melbourne, Australia), pRIT5 (Pharmacia, Piscataway, NJ) and pMAL (New England Biolabs, Beverly, MA), that fuse glutathione S-transferase, protein A, or maltose E binding protein, respectively, to a desired target recombinant protein.

Epitope tags are short peptide sequences that are recognized by epitope 10 specific antibodies. A fusion protein comprising a recombinant protein and an epitope tag can be simply and easily purified using an antibody bound to a chromatography resin, for example. The presence of the epitope tag furthermore allows the recombinant protein to be detected in subsequent assays, such as Western blots, without having to produce an antibody specific for the recombinant protein itself. 15 Examples of commonly used epitope tags include V5, glutathione-S-transferase (GST), hemagglutinin (HA), the peptide Phe-His-His-Thr-Thr, chitin binding domain, and the like.

Preferred gene expression control regions for use in avian cells include, but are not limited to, avian specific promoters such as the chicken lysozyme, ovalbumin, or 20 ovomucoid promoters, and the like. Particularly useful are tissue-specific promoters such as avian oviduct promoters that allow for expression and delivery of a heterologous polypeptide to an egg white.

Viral promoters serve the same function as bacterial or eukaryotic promoters and either provide a specific RNA polymerase in trans (bacteriophage T7) or recruit 25 cellular factors and RNA polymerase (SV40, RSV, CMV). Viral promoters may be preferred as they are generally particularly strong promoters. A preferred promoter for use in avian cells is the RSV promoter.

Selection markers are valuable elements in expression vectors as they provide a means to select for growth of only those cells that contain a vector. Common 30 selectable marker genes include those for resistance to antibiotics such as ampicillin, puromycin, tetracycline, kanamycin, bleomycin, streptomycin, hygromycin, neomycin,

ZEOCIN™, and the like.

Another element useful in an expression vector is an origin of replication. Replication origins are unique DNA segments that contain multiple short repeated sequences that are recognized by multimeric origin-binding proteins and that play a 5 key role in assembling DNA replication enzymes at the origin site. Suitable origins of replication for use in expression vectors employed herein include *E. coli* oriC, colE1 plasmid origin, and the like.

A further useful element in an expression vector is a multiple cloning site or polylinker. Synthetic DNA encoding a series of restriction endonuclease recognition 10 sites is inserted into a vector, for example, downstream of the promoter element. These sites are engineered for convenient cloning of DNA into the vector at a specific position.

Elements such as the foregoing can be combined to produce expression vectors suitable for use in the methods of the invention. Those of skill in the art will be able 15 to select and combine the elements suitable for use in their particular system in view of the teachings of the present specification.

(b) *Genetically modified avian and artificial chromosomes:* The present invention further provides modified chromosomes, either isolated avian or artificial 20 chromosomes, are useful vectors to shuttle transgenes or gene clusters into the avian genome. By delivering the modified or artificial chromosome to an isolated recipient cell, the target cell, and progeny thereof, become trisomic. Preferably, an additional or trisomic chromosome will not affect the subsequent development of the recipient cell and/or an embryo, nor interfere with the reproductive capacity of an adult bird 25 developed from such cells or embryos. The chromosome also should be stable within chicken cells. An effective method is also required to isolate a population of chromosomes for delivery into chicken embryos or early cells.

A number of artificial chromosomes are useful in the methods of the invention, including, for instance, a human chromosome modified to work as an artificial chromosome in a heterologous species as described, for example, for mice 30 (Tomizuka *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 97: 722-727 (2000); for cattle (Kuroiwa *et al.*, *Nat. Biotechnol.* 20: 889-894 (2002); a mammalian artificial

chromosome used in mice (Co *et al.*, *Chromosome Res.* 8: 183-191 (2000), or in viable triploid chickens (Thorne *et al.*, *Cytogenet. Cell Genet.* 57: 206-210 (1991) and Thorne, *et al.*, *J. Hered.* 88: 495-498 (1997). Chickens that are trisomic for microchromosome 16 have been described (Miller *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 93: 3958-3962 (1996); Muscarella *et al.*, *J. Cell Biol.* 101: 1749-1756 (1985). In these cases, triploidy and trisomy occurred naturally, and illustrate that an extra copy of one or more of the chicken chromosomes is compatible with normal development and reproductive capacity.

A useful chromosome isolation protocol can comprise the steps of inserting a lac-operator sequence (Robinett *et al.* *J. Cell Biol.* 135: 1685-1700 (1996) into an isolated chromosome and, optionally, inserting a desired transgene sequence within the same chromosome. Preferably, the lac operator region is a concatamer of a plurality of lac operators for the binding of multiple lac repressor molecules. Insertion can be accomplished, for instance, by identifying a region of known nucleotide sequence associated with a particular avian chromosome. A recombinant DNA molecule may be constructed that comprises the identified region, a recombination site such as attB or attP and a lac-operator concatamer. The recombinant molecule is delivered to an isolated avian cell, preferably, but not limited to, chicken DT40 cells that have elevated homologous recombination activity compared to other avian cell lines, whereupon homologous recombination will integrate the heterologous recombination site and the lac-operator concatamer into the targeted chromosome as shown in the schema illustrated in Fig. 20. A tag-polypeptide comprising a label domain and a lac repressor domain is also delivered to the cell, preferably by expression from a suitable expression vector. The nucleotide sequence coding for a GFP-lac-repressor fusion protein (Robinett *et al.*, *J. Cell Biol.* 135: 1685-1700 (1996)) may be inserted into the same chromosome as the lac-operator insert. The lac repressor sequence, however, can also be within a different chromosome. An inducible promoter may also be used to allow the expression of the GFP-lac-repressor only after chromosome is to be isolated.

Induced expression of the GFP-lac-repressor fusion protein will result in specific binding of the tag fusion polypeptide to the lac-operator sequence for

identification and isolation of the genetically modified chromosome. The tagged mitotic chromosome can be isolated using, for instance, flow cytometry as described in de Jong *et al.* *Cytometry* 35: 129-133 (1999) and Griffin *et al.* *Cytogenet. Cell Genet.* 87: 278-281 (1999).

5 A tagged chromosome can also be isolated using microcell technology requiring treatment of cells with the mitotic inhibitor colcemid to induce the formation of micronuclei containing intact isolated chromosomes within the cell. Final separation of the micronuclei is then accomplished by centrifugation in cytochalasin as described by Killary & Fournier in *Methods Enzymol.* 254: 133-152  
10 (1995). Further purification of microcells containing only the desired tagged chromosome could be done by flow cytometry. It is contemplated, however, that alternative methods to isolate the mitotic chromosomes or microcells, including mechanical isolation or the use of laser scissors and tweezers, and the like.

15 Delivery of a Site-Specific Nucleic Acid to a Recipient Avian Cell or Embryo.

(a) *Delivery of polynucleotide constructs.*

Most non-viral methods of gene transfer rely on normal mechanisms used by eukaryotic cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral gene delivery systems of the present invention rely 20 on endocytic pathways for the uptake of the subject transcriptional regulatory region and operably linked polypeptide-encoding nucleic acid by the targeted cell. Exemplary gene delivery systems of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes. Modified chromosomes as described above may be delivered to isolated avian embryonic cells for subsequent 25 introduction to an embryo.

In a representative embodiment, a nucleic acid molecule can be entrapped in liposomes bearing positive charges on their surface (e.g., lipofectins) and (optionally) which are tagged with antibodies against cell surface antigens of the target tissue (Mizuno *et al.*, 1992, *NO Shinkei Geka* 20: 547-551; PCT publication WO91/06309; 30 Japanese patent application 1047381; and European patent publication EP-A-43075, all of which are incorporated herein by reference in their entireties).

In similar fashion, the gene delivery system can comprise an antibody or cell surface ligand that is cross-linked with a gene binding agent such as polylysine (see, for example, PCT publications WO93/04701, WO92/22635, WO92/20316, WO92/19749, and WO92/06180, all of which are incorporated herein by reference in their entireties). It will also be appreciated that effective delivery of the subject nucleic acid constructs via receptor-mediated endocytosis can be improved using agents which enhance escape of genes from the endosomal structures. For instance, whole adenovirus or fusogenic peptides of the influenza HA gene product can be used as part of the delivery system to induce efficient disruption of DNA-containing endosomes (Mulligan *et al.*, 1993, *Science* 260:926; Wagner *et al.*, 1992, *Proc. Natl. Acad. Sci.* 89:7934-7938; and Christiano *et al.*, 1993, *Proc. Natl. Acad. Sci.* 90:2122-2126, all of which are incorporated herein by reference in their entireties). It is further contemplated that a recombinant nucleic acid molecule of the present invention may be delivered to a target host cell by other non-viral methods including by gene gun, microinjection, sperm-mediated transfer, or the like.

In yet another embodiment of the invention, an expression vector that comprises a heterologous attB recombination site and a region encoding a polypeptide deposited into an egg white are delivered to oviduct cells by in vivo electroporation. In this method, the luminal surface of an avian oviduct is surgically exposed. A buffered solution of the expression vector and a source of integrase activity such as a second expression vector expressing integrase (for example pCMV-int) is deposited on the luminal surface. Electroporation electrodes are then positioned on either side of the oviduct wall, the luminal electrode contacting the expression vector solution. After electroporation, the surgical incisions are closed. The electroporation will deliver the expression vectors to some, if not all, treated recipient oviduct cells to create a tissue-specific chimeric animal. Expression of the integrase allows for the integration of the heterologous polynucleotide into the nuclear genomes of recipient oviduct cells. While this method may be used with any bird, a preferred recipient is a chicken due to the size of the oviduct. More preferred is a transgenic bird that has a transgenic attP recombinant site in the nuclear genomes of recipient oviduct cells, thus increasing the efficiency of integration of the expression vector.

The attB/P integrase system is preferred in the in vivo electroporation method to allow the formation of stable genetically transformed oviduct cells that otherwise progressively lose the heterologous expression vector.

The stably modified oviduct cells will express the heterologous polynucleotide 5 and deposit the resulting polypeptide into the egg white of a laid egg. For this purpose, the expression vector will further comprise an oviduct-specific promoter such as ovalbumin or ovomucoid operably linked to the desired heterologous polynucleotide.

*(b) Delivery of chromosomes to avian cells.*

10 Another aspect of the invention is the generation of a trisomic avian cell comprising a genetically modified extra chromosome. The extra chromosome may be an artificial chromosome or an isolated avian chromosome that has been genetically modified. Introduction of the extra chromosome to an avian cell will generate a trisomic cell with  $2n+1$  chromosomes, where  $n$  is the haploid number of chromosomes 15 of a normal avian cell.

Delivery of an isolated chromosome into an isolated avian cell or embryo can be accomplished in several ways. Isolated mitotic chromosomes or a micronucleus containing an interphase chromosome can be injected into early stage I embryos by cytoplasmic injection. The injected zygote would then be surgically transferred to a 20 recipient hen for the production and laying of a hard shell egg. This hard shell egg would then be incubated until hatching of a chick.

Isolated microcells can be fused to primordial germ cells (PGCs) isolated from the blood stream of late stage 15 embryos as described by Killary & Fournier in *Methods Enzymol.* 254: 133-152 (1995). The PGC/microcell hybrids can then be 25 transplanted into the blood stream of a recipient embryo to produce germline chimeric chickens. (See Naito *et al.*, *Mol. Reprod. Dev.* 39: 153-161 (1994)). The manipulated eggs would then be incubated until hatching of the bird.

Blastodermal cells isolated from stage X embryos can be transfected with isolated mitotic chromosomes. Following *in vitro* transfection, the cells are 30 transplanted back into stage X embryos as described, for example, in Etches *et al.*, *Poult. Sci.*, 72: 882-829 (1993), and the manipulated eggs are incubated to hatching.

Stage X blastodermal cells can also be fused with isolated microcells and then transplanted back into to stage X embryos or fused to somatic cells to be used as nuclear donors for nuclear transfer as described by Kuroiwa *et al.*, *Nat. Biotechnol.* 20: 889-894 (2002).

5 Chromosomal vectors, as described above, may be delivered to a recipient avian cell by, for example, microinjection, liposomal delivery or microcell fusion.

*Delivering a Source of Integrase Activity to an Avian Cell*

In the methods of the invention, a site-specific integrase is introduced into an 10 avian cell whose genome is to be modified. Methods of introducing functional proteins into cells are well known in the art. Introduction of purified integrase protein can ensure a transient presence of the protein and its activity. Thus, the lack of permanence associated with most expression vectors is not expected to be detrimental.

The integrase used in the practice of the present invention can be introduced 15 into a target cell before, concurrently with, or after the introduction of a site-specific vector. The integrase can be directly introduced into a cell as a protein, for example, by using liposomes, coated particles, or microinjection, or into the blastodermal layer of an early stage avian embryo by microinjection. A source of the integrase can also be delivered to an avian cell by introducing to the cell an mRNA encoding the 20 integrase and which can be expressed in the recipient cell as an integrase polypeptide. Alternately, a DNA molecule encoding the integrase can be introduced into the cell using a suitable expression vector.

The present invention provides novel nucleic acid vectors and methods of use that allow the phiC31 integrase to efficiently integrate a heterologous nucleic acid into 25 an avian genome. A novel finding is that the phiC31 integrase is remarkably efficient in avian cells and increases the rate of integration of heterologous nucleic acid at least 30-fold over that of random integration. Furthermore, the phiC31 integrase works equally well at 37°C and 41°C, indicating that it will function in the environment of the developing avian embryo, as shown in Example 1.

30 The site-specific vector components described above are useful in the construction of expression cassettes containing sequences encoding an integrase. One

integrase-expressing vector useful in the methods of the invention is pCMV-C31int (SEQ ID NO: 1 as shown in Fig. 9) where the phiC31 integrase is encoded by a region under the expression control of the strong CMV promoter. Another preferred promoter generally useful in avian cells is the RSV promoter as used in SEQ ID NO: 9 5 shown in Fig. 17. Expression of the integrase is typically desired to be transient. Accordingly, vectors providing transient expression of the integrase are preferred. However, expression of the integrase can be regulated in other ways, for example, by placing the expression of the integrase under the control of a regulatable promoter (i.e., a promoter whose expression can be selectively induced or repressed).

10

Transgenic Avian Cells.

Another aspect of the present invention is an avian cell genetically modified with a transgene vector according to the present invention and described above. For example, in one embodiment, the transformed cell can be a chicken early stage 15 blastodermal cell or a genetically transformed cell line, including a sustainable cell line. The transfected cell according to the present invention may comprise a transgene stably integrated into the nuclear genome of the recipient cell, thereby replicating with the cell so that each progeny cell receives a copy of the transfected nucleic acid. A particularly useful cell line for the delivery and integration of a transgene comprises a 20 heterologous attP site that can increase the efficiency of integration of a polynucleotide by phiC31 integrase and, optionally, a region for expressing the integrase.

A retroviral vector can be used to deliver the att site into the avian genome since an attP or attB site is less than 300 bp. For example, the attP site can be inserted 25 into the NLB retroviral vector, which is based on the avian leukosis virus genome. A lentiviral vector is a particularly suitable vector because lentiviral vectors can transduce non-dividing cells, so that a higher percentage of cells will have an integrated attP site.

The lacZ region of NLB is replaced by the attP sequence. A producer cell line 30 would be created by transformation of, for example, the Isolde cell line capable of producing a packaged recombinant NLB-attP virus pseudo-typed with the envA

envelope protein. Supernatant from the Isolde NLB-attP line is concentrated by centrifugation to produce high titer preparations of the retroviral vector that can then be used to deliver the attP site to the genome of an avian cell, as described in Example 9 below.

5 An attP-containing line of transgenic birds are a source of attP transgenic embryos and embryonic cells. Fertile zygotes and oocytes bearing a heterologous attP site in either the maternal, paternal, or both, genomes can be used for transgenic insertion of a desired heterologous polynucleotide. A transgene vector bearing an attB site, for example, would be injected into the cytoplasm along with either an integrase 10 expression plasmid, mRNA encoding the integrase or the purified integrase protein. The oocyte or zygote is then cultured to hatch by *ex ovo* methods or reintroduced into a recipient hen such that the hen lays a hard shell egg the next day containing the injected egg.

In another example, fertile stage VII-XII embryos hemizygous or homozygous 15 for the heterologous attP sequence, are used as a source of blastodermal cells. The cells are harvested and then transfected with a transgene vector bearing an attB site along with a source of integrase. The transfected cells are then injected into the subgerminal cavity of windowed fertile eggs. The chicks that hatch will bear the transgene integrated into the attP site in a percentage of their somatic and germ cells. 20 To obtain fully transgenic birds, chicks are raised to sexual maturity and those that are positive for the transgene in their semen are bred to non-transgenic mates.

In various embodiments, the genetically engineered cells of the invention may contain an integrase specifically recognizing recombination sites and which is introduced into genetically engineered cells containing a nucleic acid construct of the 25 invention under conditions such that the nucleic acid sequence(s) of interest will be inserted into the nuclear genome. Methods for introducing such an integrase into a cell are described above.

In some embodiments, the site-specific integrase is introduced into the cell as a polypeptide. In alternative embodiments, the site-specific integrase is introduced into 30 the transgenic cell as a polynucleotide encoding the integrase, such as an expression

cassette optionally carried on a transient expression vector, and comprising a polynucleotide encoding the recombinase.

In one embodiment, the invention is directed to methods of using a vector for site-specific integration of a heterologous nucleotide sequence into the genome of an avian cell, the vector comprising a circular backbone vector, a polynucleotide of interest operably linked to a promoter, and a first recombination site, wherein the genome of the cell comprises a second recombination site and recombination between the first and second recombination sites is facilitated by phiC31 integrase. In certain embodiments, the integrase facilitates recombination between a bacterial genomic recombination site (attB) and a phage genomic recombination site (attP).

In another embodiment, the invention is directed to an avian cell having a transformed genome comprising an integrated heterologous polynucleotide of interest whose integration, mediated by phiC31 integrase, was into a recombination site native to the avian cell genome and the integration created a recombination-product site comprising the polynucleotide sequence. In yet another embodiment, integration of the polynucleotide was into a recombination site not native to the avian cell genome, but instead into a heterologous recombination site engineered into the avian cell genome.

In further embodiments, the invention is directed to transgenic birds comprising a modified cell and progeny thereof as described above, as well as methods of producing the same.

Cells genetically modified to carry a heterologous attB or attP site by the methods of the present invention can be maintained under conditions that, for example, keep them alive but do not promote growth, promote growth of the cells, and/or cause the cells to differentiate or dedifferentiate. Cell culture conditions may be permissive for the action of the integrase in the cells, although regulation of the activity of the integrase may also be modulated by culture conditions (e.g., raising or lowering the temperature at which the cells are cultured).

One aspect of the invention is a method for generating a genetically modified avian cell, and progeny thereof, using a tagged chromosome, the method comprising the steps of providing an isolated modified chromosome comprising a lac operator

region and a first recombination site, delivering the modified chromosome to a avian cell, thereby generating a trisomic avian cell, delivering to the avian cell a source of a tagged polypeptide comprising a fluorescent domain and a lac repressor domain, delivering a source of integrase activity to the avian cell, delivering a polynucleotide 5 comprising a second recombination site and a region encoding a polypeptide to the avian cell, maintaining the avian cell under conditions suitable for the integrase to mediate recombination between the first and second recombination sites, thereby integrating the polynucleotide into the modified chromosome and generating a genetically modified avian cell, expressing the tag polypeptide by the avian cell, 10 allowing the tag polypeptide to bind to the modified chromosome so as to label the modified chromosome, and isolating the modified chromosome by selecting modified chromosomes having a tag polypeptide bound thereto.

In one embodiment of the invention, the second avian cell is selected from the group consisting of a stage VII-XII blastodermal cell, a stage I embryo, a stage X 15 embryo; an isolated primordial germ cell, an isolated non-embryonic cell, and an oviduct cell.

In various embodiments, the isolated modified chromosome is an avian chromosome or an artificial chromosome.

In other embodiments of the invention, the step of providing an isolated 20 modified chromosome comprising a lac operator region and a first recombination site comprises the steps of generating a trisomic avian cell by delivering to an isolated avian cell an isolated chromosome and a polynucleotide comprising a lac operator and a second recombination site, maintaining the trisomic cell under conditions whereby the heterologous polynucleotide is integrated into the chromosome by homologous 25 recombination, delivering to the avian cell a source of a tag polypeptide to label the chromosome, and isolating the labeled chromosome.

In one embodiment of the invention, the lac operator region is a concatamer of lac operators. In other embodiments of the invention, the tag polypeptide is expressed from an expression vector.

30 In one embodiment of the invention, the tag polypeptide is microinjected into the cell. In various embodiments of the invention, the method of delivery of a

chromosome to an avian cell is selected from the group consisting of liposome delivery, microinjection, microcell, electroporation and gene gun delivery, or a combination thereof.

In embodiments of the invention, the fluorescent domain of the tag  
5 polypeptide is GFP.

In another embodiment of the invention, the method further comprises the step of delivering the second avian cell to an avian embryo. The embryo may be maintained under conditions suitable for hatching as a chick.

In one embodiment of the invention, the second avian cell is maintained under  
10 conditions suitable for the proliferation of the cell, and progeny thereof.

In various embodiments of the invention, the source of integrase activity is delivered to a first avian cell as a polypeptide or expressed from a polynucleotide, said polynucleotide being selected from an mRNA and an expression vector.

In one embodiment of the invention, the tag polypeptide activity is delivered  
15 to the avian cell as a polypeptide or expressed from a polynucleotide operably linked to a promoter. In another embodiment of the invention, the promoter is an inducible promoter. In yet another embodiment of the invention, the integrase is phiC31 integrase and in various embodiments of the invention, the first and second recombination sites are selected from an attB and an attP site, but wherein the first and  
20 second sites are not identical.

Expression of Heterologous Proteins by Site-Specific Genetic Transformation of Avian Cells.

Another aspect of the present invention is a method of expressing a  
25 heterologous polypeptide in an avian cell by stably transfecting a cell by using site-specific integrase-mediation and a recombinant nucleic acid molecule, as described above, and culturing the transfected cell under conditions suitable for expression of the heterologous polypeptide under the control of the avian transcriptional regulatory region.

30 The protein of the present invention may be produced in purified form by any known conventional techniques. For example, chicken cells, an egg or an egg white

may be homogenized and centrifuged. The supernatant may then be subjected to sequential ammonium sulfate precipitation and heat treatment. The fraction containing the protein of the present invention is subjected to gel filtration in an appropriately sized dextran or polyacrylamide column to separate the proteins. If 5 necessary, the protein fraction may be further purified by HPLC or other methods well known in the art of protein purification.

The methods of the invention are useful for expressing nucleic acid sequences that are optimized for expression in avian cells and which encode desired polypeptides or derivatives and fragments thereof. Derivatives include, for instance, polypeptides 10 with conservative amino acid replacements, that is, those within a family of amino acids that are related in their side chains (commonly known as acidic, basic, nonpolar, and uncharged polar amino acids). Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids and other groupings are known in the art (see, for example, "Biochemistry", 2nd ed, L. Stryer, ed., W.H. Freeman & 15 Co., 1981). Peptides in which more than one replacement has taken place can readily be tested for activity in the same manner as derivatives with a single replacement, using conventional polypeptide activity assays (e.g. for enzymatic or ligand binding activities).

Regarding codon optimization, if the recombinant nucleic acid molecules are 20 transfected into a recipient chicken cell, the sequence of the nucleic acid insert to be expressed can be optimized for chicken codon usage. This may be determined from the codon usage of at least one, and preferably more than one, protein expressed in a chicken cell according to well known principles. For example, in the chicken the codon usage could be determined from the nucleic acid sequences encoding the 25 proteins such as lysozyme, ovalbumin, ovomucin and ovotransferrin of chicken. Optimization of the sequence for codon usage can elevate the level of translation in avian eggs.

The present invention provides methods for the production of a protein by an avian cell comprising the steps of maintaining an avian cell, transfecting with a first 30 expression vector and, optionally, a second expression vector, under conditions suitable for proliferation and/or gene expression and such that an integrase will

mediate site specific recombination at att sites. The expression vectors may each have a transcription unit comprising a nucleotide sequence encoding a heterologous polypeptide, wherein one polypeptide is an integrase, a transcription promoter, and a transcriptional terminator. The cells may then be maintained under conditions for the 5 expression and production of the desired heterologous polypeptide(s).

The present invention further relates to methods for gene expression by avian cells from nucleic acid vectors, and transgenes derived therefrom, that include more than one polypeptide-encoding region wherein, for example, a first polypeptide-encoding region can be operatively linked to an avian promoter and a second 10 polypeptide-encoding region is operatively linked to an Internal Ribosome Entry Sequence (IRES). It is contemplated that the first polypeptide-encoding region, the IRES and the second polypeptide-encoding region of a recombinant DNA of the present invention may be arranged linearly, with the IRES operably positioned immediately 5' of the second polypeptide-encoding region. This nucleic acid 15 construct, when inserted into the genome of an avian cell or a bird and expressed therein, will generate individual polypeptides that may be post-translationally modified and combined in the white of a hard shell bird egg. Alternatively, the expressed polypeptides may be isolated from an avian egg and combined *in vitro*.

The invention, therefore, includes methods for producing multimeric proteins 20 including immunoglobulins, such as antibodies, and antigen binding fragments thereof. Thus, in one embodiment of the present invention, the multimeric protein is an immunoglobulin, wherein the first and second heterologous polypeptides are immunoglobulin heavy and light chains respectively. Illustrative examples of this and other aspects of the present invention for the production of heterologous multimeric 25 polypeptides in avian cells are fully disclosed in U.S. Patent Application No. 09/877,374, filed June 8, 2001, by *Rapp*, published as US-2002-0108132-A1 on August 8, 2002, and U.S. Patent Application No. 10/251,364, filed September 18, 2002, by *Rapp*, both of which are incorporated herein by reference in their entirety.

Accordingly, the invention further provides immunoglobulin and other 30 multimeric proteins that have been produced by transgenic avians of the invention.

In various embodiments, an immunoglobulin polypeptide encoded by the transcriptional unit of at least one expression vector may be an immunoglobulin heavy chain polypeptide comprising a variable region or a variant thereof, and may further comprise a D region, a J region, a C region, or a combination thereof. An 5 immunoglobulin polypeptide encoded by an expression vector may also be an immunoglobulin light chain polypeptide comprising a variable region or a variant thereof, and may further comprise a J region and a C region. The present invention also contemplates multiple immunoglobulin regions that are derived from the same animal species, or a mixture of species including, but not only, human, mouse, rat, 10 rabbit and chicken. In preferred embodiments, the antibodies are human or humanized.

In other embodiments, the immunoglobulin polypeptide encoded by at least one expression vector comprises an immunoglobulin heavy chain variable region, an immunoglobulin light chain variable region, and a linker peptide thereby forming a 15 single-chain antibody capable of selectively binding an antigen.

Examples of therapeutic antibodies that can be used in methods of the invention include but are not limited to HERCEPTIN<sup>TM</sup> (Trastuzumab) (Genentech, CA) which is a humanized anti-HER2 monoclonal antibody for the treatment of patients with metastatic breast cancer; REOPRO<sup>TM</sup> (abciximab) (Centocor) which is 20 an anti-glycoprotein IIb/IIIa receptor on the platelets for the prevention of clot formation; ZENAPAX<sup>TM</sup> (daclizumab) (Roche Pharmaceuticals, Switzerland) which is an immunosuppressive, humanized anti-CD25 monoclonal antibody for the prevention of acute renal allograft rejection; PANOREX<sup>TM</sup> which is a murine anti-17-IA cell surface antigen IgG2a antibody (Glaxo Wellcome/Centocor); BEC2 which is a 25 murine anti-idiotype (GD3 epitope) IgG antibody (ImClone System); IMC-C225 which is a chimeric anti-EGFR IgG antibody (ImClone System); VITAXIN<sup>TM</sup> which is a humanized anti- $\alpha$ V $\beta$ 3 integrin antibody (Applied Molecular Evolution/MedImmune); Campath 1H/LDP-03 which is a humanized anti CD52 IgG1 antibody (Leukosite); Smart M195 which is a humanized anti-CD33 IgG antibody 30 (Protein Design Lab/Kanebo); RITUXAN<sup>TM</sup> which is a chimeric anti-CD20 IgG1 antibody (IDEC Pharm/Genentech, Roche/Zettyaku); LYMPHOCIDE<sup>TM</sup> which is a

humanized anti-CD22 IgG antibody (Immunomedics); ICM3 is a humanized anti-ICAM3 antibody (ICOS Pharm); IDEC-114 is a primate anti-CD80 antibody (IDEC Pharm/Mitsubishi); ZEVALINT<sup>TM</sup> is a radiolabelled murine anti-CD20 antibody (IDEC/Schering AG); IDEC-131 is a humanized anti-CD40L antibody (IDEC/Eisai);  
5 IDEC-151 is a primatized anti-CD4 antibody (IDEC); IDEC-152 is a primatized anti-CD23 antibody (IDEC/Seikagaku); SMART anti-CD3 is a humanized anti-CD3 IgG (Protein Design Lab); 5G1.1 is a humanized anti-complement factor 5 (CS) antibody (Alexion Pharm); D2E7 is a humanized anti-TNF- $\alpha$  antibody (CATIBASF); CDP870 is a humanized anti-TNF- $\alpha$  Fab fragment (Celltech); IDEC-151 is a primatized anti-  
10 CD4 IgG1 antibody (IDEC Pharm/SmithKline Beecham); MDX-CD4 is a human anti-CD4 IgG antibody (Medarex/Eisai/Genmab); CDP571 is a humanized anti-TNF- $\alpha$  IgG4 antibody (Celltech); LDP-02 is a humanized anti- $\alpha$ 4 $\beta$ 7 antibody (LeukoSite/Genentech); OrthoClone OKT4A is a humanized anti-CD4 IgG antibody (Ortho Biotech); ANTOVA<sup>TM</sup> is a humanized anti-CD40L IgG antibody (Biogen);  
15 ANTEGRENT<sup>TM</sup> is a humanized anti-VLA-4 IgG antibody (Elan); and CAT-152 is a human anti-TGF- $\beta$ <sub>2</sub> antibody (Cambridge Ab Tech).

Production of Heterologous Protein by Transgenic Avians

One aspect of the present invention, therefore, concerns transgenic birds, such as chickens, comprising a recombinant nucleic acid molecule and which preferably (though optionally) express a heterologous gene in one or more cells in the animal. Suitable methods for the generation of transgenic avians having heterologous DNA incorporated therein are described, for example, in WO 99/19472 to Ivarie et al.; WO 00/11151 to Ivarie et al.; and WO 00/56932 to Harvey et al., all of which are incorporated herein by reference in their entirety.

Embodiments of the methods for the production of a heterologous polypeptide by the avian tissue such as the oviduct and the production of eggs which contain heterologous protein involve providing a suitable vector and introducing the vector into embryonic blastodermal cells together with an integrase, preferably phiC31 integrase, so that the vector can integrate into the avian genome. A subsequent step involves deriving a mature transgenic avian from the transgenic blastodermal cells

produced in the previous steps. Deriving a mature transgenic avian from the blastodermal cells optionally involves transferring the transgenic blastodermal cells to an embryo and allowing that embryo to develop fully, so that the cells become incorporated into the bird as the embryo is allowed to develop. Another alternative is 5 to transfer a transfected nucleus to an enucleated recipient cell which may then develop into a zygote and ultimately an adult bird. The resulting chick is then grown to maturity.

In an alternative embodiment, the cells of a blastodermal embryo are transfected or transduced with the vector and integrase directly within the embryo. It 10 is contemplated, for example, that the recombinant nucleic acid molecules of the present invention may be introduced into a blastodermal embryo by direct microinjection of the DNA into a stage X or earlier embryo that has been removed from the oviduct. The egg is then returned to the bird for egg white deposition, shell development and laying. The resulting embryo is allowed to develop and hatch, and 15 the chick allowed to mature.

In one embodiment, a transgenic bird of the present invention is produced by introducing into embryonic cells such as, for instance, isolated avian blastodermal cells, a nucleic acid construct comprising an attB recombination site capable of recombining with a pseudo-attP recombination site found within the nuclear genome 20 of the organism from which the cell was derived, and a nucleic acid fragment of interest, in a manner such that the nucleic acid fragment of interest is stably integrated into the nuclear genome of germ line cells of a mature bird and is inherited in normal Mendelian fashion. It is also within the scope of the invention that the targeted cells 25 for receiving the transgene have been engineered to have a heterologous attP recombination site integrated into the nuclear genome of the cells, thereby increasing the efficiency of recognition and recombination with a heterologous attB site.

In either case, the transgenic bird produced from the transgenic blastodermal cells is known as a "founder" Some founders can be chimeric or mosaic birds if, for example, microinjection does not deliver nucleic acid molecules to all of the 30 blastodermal cells of an embryo. Some founders will carry the transgene in the tubular gland cells in the magnum of their oviducts and will express the heterologous

protein encoded by the transgene in their oviducts. If the heterologous protein contains the appropriate signal sequences, it will be secreted into the lumen of the oviduct and onto the yolk of an egg.

Some founders are germ-line founders. A germ-line founder is a founder that 5 carries the transgene in genetic material of its germ-line tissue, and may also carry the transgene in oviduct magnum tubular gland cells that express the heterologous protein. Therefore, in accordance with the invention, the transgenic bird will have tubular gland cells expressing the heterologous protein and the offspring of the transgenic bird will also have oviduct magnum tubular gland cells that express the 10 selected heterologous protein. (Alternatively, the offspring express a phenotype determined by expression of the exogenous gene in a specific tissue of the avian.)

The invention can be used to express, in large yields and at low cost, a wide range of desired proteins including those used as human and animal pharmaceuticals, diagnostics, and livestock feed additives. Proteins such as growth hormones, 15 cytokines, structural proteins and enzymes including human growth hormone, interferon, lysozyme, and  $\beta$ -casein are examples of proteins which are desirably expressed in the oviduct and deposited in eggs according to the invention. Other possible proteins to be produced include, but are not limited to, albumin,  $\alpha$ -1 antitrypsin, antithrombin III, collagen, factors VIII, IX, X (and the like), fibrinogen, 20 hyaluronic acid, insulin, lactoferrin, protein C, erythropoietin (EPO), granulocyte colony-stimulating factor (G-CSF), granulocyte macrophage colony-stimulating factor (GM-CSF), tissue-type plasminogen activator (tPA), feed additive enzymes, somatotropin, and chymotrypsin. Immunoglobulins (shown, for example in Example 10 below) and genetically engineered antibodies, including immunotoxins which bind 25 to surface antigens on human tumor cells and destroy them, can also be expressed for use as pharmaceuticals or diagnostics.

In various embodiments of the transgenic bird of the present invention, the expression of the transgene may be restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, *trans*-acting factors acting on the 30 transcriptional regulatory region operably linked to the polypeptide-encoding region of interest of the present invention and which control gene expression in the desired

pattern. Tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory 5 sequences.

The stably modified oviduct cells will express the heterologous polynucleotide and deposit the resulting polypeptide into the egg white of a laid egg. For this purpose, the expression vector will further comprise an oviduct-specific promoter such as ovalbumin or ovomucoid operably linked to the desired heterologous 10 polynucleotide.

Another aspect of the present invention provides a method for the production in an avian of an heterologous protein capable of forming an antibody suitable for selectively binding an antigen. This method comprises a step of producing a transgenic avian incorporating at least one transgene, the transgene encoding at least 15 one heterologous polypeptide selected from an immunoglobulin heavy chain variable region, an immunoglobulin heavy chain comprising a variable region and a constant region, an immunoglobulin light chain variable region, an immunoglobulin light chain comprising a variable region and a constant region, and a single-chain antibody comprising two peptide-linked immunoglobulin variable regions.

20 In one embodiment of this method, the isolated heterologous protein is an antibody capable of selectively binding to an antigen and which may be generated by combining at least one immunoglobulin heavy chain variable region and at least one immunoglobulin light chain variable region, preferably cross-linked by at least one disulfide bridge. The combination of the two variable regions generates a binding site 25 that binds an antigen using methods for antibody reconstitution that are well known in the art.

The present invention also encompasses immunoglobulin heavy and light chains, or variants or derivatives thereof, to be expressed in separate transgenic avians, and thereafter isolated from separate media including serum or eggs, each 30 isolate comprising one or more distinct species of immunoglobulin polypeptide. The method may further comprise the step of combining a plurality of isolated

heterologous immunoglobulin polypeptides, thereby producing an antibody capable of selectively binding to an antigen. In this embodiment, for instance, two or more individual transgenic avians may be generated wherein one transgenic produces serum or eggs having an immunoglobulin heavy chain variable region, or a polypeptide 5 comprising such, expressed therein. A second transgenic animal, having a second transgene, produces serum or eggs having an immunoglobulin light chain variable region, or a polypeptide comprising such, expressed therein. The polypeptides from two or more transgenic animals may be isolated from their respective sera and eggs and combined in vitro to generate a binding site capable of binding an antigen.

10

The present invention is further illustrated by the following examples, which are provided by way of illustration and should not be construed as limiting. The contents of all references, published patents and patents cited throughout the present application are hereby incorporated by reference in their entireties.

15

It will be apparent to those skilled in the art that various modifications, combinations, additions, deletions and variations can be made in the present invention without departing from the scope or spirit of the invention. For instance, features illustrated or described as part of one embodiment can be used in another embodiment to yield a still further embodiment. It is intended that the present invention covers 20 such modifications, combinations, additions, deletions and variations as come within the scope of the appended claims and their equivalents.

20

**Example 1: Phage phiC31 integrase functions in avian cells.**

(a) A luciferase vector bearing either an attB (SEQ ID NO: 2 shown in Fig. 10) or attP 25 (SEQ ID NO: 3 shown in Fig. 11) site was co-transfected with an integrase expression vector CMV-C31int (SEQ ID NO: 1) into DF-1 cells, a chicken fibroblast cell line. The cells were passaged several times and the luciferase levels were assayed at each passage.

Cells were passaged every 3-4 days and one third of the cells were harvested 30 and assayed for luciferase. The expression of luciferase was plotted as a percentage of

the expression measured 4 days after transfection. A luciferase expression vector bearing an attP site as a control was also included.

As can be seen in Fig. 2, in the absence of integrase, luciferase expression from a vector bearing attP or attB decreased to very low levels after several days. 5 However, luciferase levels were persistent when the luciferase vector bearing attB was co-transfected with the integrase expression vector, indicating that the luciferase vector had stably integrated into the avian genome.

(b) A drug-resistance colony formation assay was used to quantitate integration efficiency. The puromycin resistance expression vector pCMV-pur was outfitted with 10 an attB (SEQ ID NO: 4 shown in Fig. 12) or an attP (SEQ ID NO: 5 shown in Fig. 13) sites. Puromycin resistance vectors bearing attB sites were cotransfected with phiC31 integrase or a control vector into DF-1 cells. One day after transfection, puromycin was added. Puromycin resistant colonies were counted 12 days post-transfection.

In the absence of co-transfected integrase expression, few DF-1 cell colonies 15 were observed after survival selection. When integrase was co-expressed, multiple DF-1 cell colonies were observed, as shown in Fig. 3. Similar to the luciferase expression experiment, the attB sequence (but not the attP sequence) was able to facilitate integration of the plasmid into the genome. Fig. 3 also shows that phiC31 integrase functions at both 37° Celsius and 41° Celsius. Integrase also functions in 20 quail cells using the puromycin resistance assay, as shown in Fig. 4.

(c) The CMV-pur-attB vector (SEQ ID NO: 4) was also cotransfected with an enhanced green fluorescent protein (EGFP) expression vector bearing an attB site (SEQ ID NO: 6 shown in Fig. 14) into DF-1 cells and the phiC31 integrase expression vector CMV-C31int (SEQ ID NO: 1). After puromycin selection for 12 days, the 25 colonies were viewed with UV light to determine the percentage of cells that expressed EGFP. Approximately 20% of puromycin resistant colonies expressed EGFP in all of the cells of the colony, as shown in Fig. 5, indicating that the integrase can mediate multiple integrations per cell.

(d) PhiC31 integrase promoted the integration of large transgenes into avian 30 cells. A puromycin expression cassette comprising a CMV promoter, puromycin resistance gene, polyadenylation sequence and the attB sequence was inserted into a

vector containing a 12.0 kb lysozyme promoter and the human interferon  $\alpha$ 2b gene (SEQ ID NO: 7 shown in Fig. 15) and into a vector containing a 10.0 kb ovomucoid promoter and the human interferon  $\alpha$ 2b gene (SEQ ID NO: 8) as shown in Fig. 16.

DF-1 cells were transfected with donor plasmids of varying lengths bearing a puromycin resistance gene and an attB sequence in the absence or presence of an integrase expression plasmid. Puromycin was added to the culture media to kill those cells which did not contain a stably integrated copy of the puromycin resistance gene. Cells with an integrated gene formed colonies in the presence of puromycin in 7-12 days. The colonies were visualized by staining with methylene blue and the entire 60 mm culture dish was imaged.

PhiC31 integrase mediated the efficient integration of both vectors as shown in Fig. 7.

#### Example 2: Cell culture methods.

DF-1 cells were cultured in DMEM with high glucose, 10% fetal bovine serum, 2 mM L-glutamine, 100 units/ml penicillin and 100  $\mu$ g/ml streptomycin at 37° Celsius and 5% CO<sub>2</sub>. A separate population of DF-1 cells was grown at 41° Celsius. These cells were adapted to the higher temperature for one week before they were used for experiments.

Quail QT6 cells were cultured in F10 medium (Gibco) with 5% newborn calf serum, 1% chicken serum heat inactivated (at 55° Celsius for 45 mins), 10 units/ ml penicillin and 10  $\mu$ g/ml streptomycin at 37° Celsius and 5% CO<sub>2</sub>.

#### Example 3: Selection and Assay Methods

(a) *Puromycin selection assay*: About  $0.8 \times 10^6$  DF-1 (chicken) or QT6 (quail) cells were plated in 60 mm dishes. The next day, the cells were transfected as follows:

10 to 50 ng of a donor plasmid and 1 to 10  $\mu$ g of an Integrase-expressing plasmid DNA were mixed with 150  $\mu$ l of OptiMEM. 15  $\mu$ l of DMRIE-C was mixed with 150  $\mu$ l of OptiMEM in a separate tube, and the mixtures combined and incubated for 15 mins. at room temperature.

While the liposome/DNA complexes were forming, the cells were washed with OptiMEM and 2.5 ml of OptiMEM was added. After 15 minutes, 300  $\mu$ l of the DNA-lipid mixture was added dropwise to the 2.5 ml of OptiMEM covering the cell layers. The cells were incubated for 4-5 hours at either 37° Celsius or 41° Celsius, 5% 5  $\text{CO}_2$ . The transfection mix was replaced with 3 mls of culture media. The next day, puromycin was added to the media at a final concentration of 1 ug/ml, and the media replaced every 2 to 4 days. Puromycin resistant colonies were counted or imaged 10-12 days after the addition of puromycin.

10 (b) *Luciferase assay*: Chicken DF-1 or quail QT6 cells ( $0.8 \times 10^6$ ) were plated in 60 mm dishes. Cells were transfected as described above. The cells from a plate were transferred to a new 100 mm plate when the plate became confluent, typically on day 3-4, and re-passaged every 3-4 days.

15 At each time point, one-third of the cells from a plate were replated, and one-third were harvested for the luciferase assay. The cells were pelleted in an eppendorf tube and frozen at -70°C.

The cell pellet was lysed in 200  $\mu$ l of lysis buffer (25 mM Tris-acetate, pH7.8, 2mM EDTA, 0.5% Triton X-100, 5% glycerol). Sample (5 $\mu$ l) was assayed using the Promega BrightGlo reagent system.

20 (c) *Visualization of EGFP*: EGFP expression was visualized with an inverted microscope with FITC illumination [Olympus IX70, 100 W mercury lamp, HQ-FITC Band Pass Emission filter cube, exciter 480/40 nm, emission 535/50 nm, 20X phase contrast objective (total magnification was 2.5 x 10 x 20)].

25 (d) *Staining of cell colonies*: After colonies had formed, typically after 7-12 days of culture in puromycin medium, the cells were fixed in 2% formaldehyde, 0.2% glutaraldehyde for 15 mins, and stained in 0.2% methylene blue for 30 mins. followed by several washes with water. The plates were imaged using a standard CCD camera in visible light.

**Example 4: Generation of genetically transferred avian cells.**

30 Avian stage X blastodermal cells are used as the cellular vector for the transgenes. Stage X embryos are collected and the cells dispersed and mixed with

plasmid DNA. The transgenes are then introduced to blastodermal cells via electroporation. The cells are immediately injected back into recipient embryos.

The cells are not cultured for any time period to ensure that they remain capable of contributing to the germline of resulting chimeric embryos. However, 5 because there is no culture step, cells that bear the transgene cannot be identified. Typically, only a small percentage of cells introduced to an embryo will bear a stably integrated transgene (0.01 to 1%). To increase the percentage of cells bearing a transgene, therefore, the transgene vector bears an attB site and is co-electroporated with a vector bearing the CMV promoter driving expression of the phiC31 transgene 10 (CMV-C31int (SEQ ID NO: 1)). The integrase then drives integration of the transgene vector into the nuclear genome of the avian cell and increases the percentage of cells bearing a stable transgene.

*(a) Preparation of avian stage X blastodermal cells:*

- i) Collect fertilized eggs from Barred Rock or White leghorn chickens 15 (*Gallus gallus*) or quail (*Japonica coturnix*) within 48 hrs. of laying;
- ii) Use 70% ethanol to clean the shells;
- iii) Crack the shells and open the eggs;
- iv) Remove egg whites by transferring yolks to opposite halves of shells, repeating to remove most of the egg whites;
- 20 v) Put egg yolks with embryo discs facing up into a 10cm petri dish;
- vi) Use an absorbent tissue to gently remove egg white from the embryo discs;
- vii) Place a Whatman filter paper 1 ring over the embryos;
- viii) Use scissors to cut the membranes along the outside edge of the paper ring while gently lifting the ring/embryos with a pair of tweezers;
- 25 ix) Insert the paper ring with the embryos at a 45 degrees angle into a petri dish containing PBS-G solution at room temperature;
- x) After ten embryo discs are collected, gently wash the yolks from the blastoderm discs using a Pasteur pipette under a stereo microscope;
- xi) Cut the discs by a hair ring cutter (a short piece of human hair is bent into a small loop and fastened to the narrow end of a Pasteur pipette with Parafilm);

- xii) Transfer the discs to a 15 ml sterile centrifuge tube on ice;
- xiii) Place 10 to 15 embryos per tube and allow to settle to the bottom (about 5 mins.);
- xiv) Aspirate the supernatant from the tube;
- 5 xv) Add 5 mls of ice-cold PBS without  $\text{Ca}^{++}$  and  $\text{Mg}^{++}$ , and gently pipette 4 to 5 times using a 5 mls pipette;
- xvi) Incubate in ice for 5-7 mins. to allow the blastoderms to settle, and aspirate the supernatant;
- xvii) Add 3 mls of ice cold 0.05% trypsin/0.02% ETDA to each tube and gently pipette 3 to 5 times using a 5 ml pipette;
- 10 xviii) Put the tube in ice for 5 mins. and then flick the tube by finger 40 times. Repeat;
- xix) Add 0.5 mls FBS and 3-5 mls BDC medium to each tube and gently pipette 5-7 times using a 5 ml pipette;
- 15 xx) Spin at 500 rpm (RCF 57 x g) at 4° Celsius for 5 mins;
- xxi) Remove the supernatant and add 2 mls ice cold BDC medium into each tube; and
- xxii) Resuspend the cells by gently pipetting 20-25 times; and
- xxiii) Determine the cell titer by hemacytometer and ensure that about 95% of all
- 20 BDCs are single cells, and not clumped.

*(b) Transfection of linearized plasmids into blastodermal cells by small scale electroporation:*

- i) Centrifuge the blastodermal cell suspension from step (xxiii) above at RCF 57 x g, 4° Celsius, for 5 mins;
- 25 ii) Resuspend cells to a density of  $1\text{-}3 \times 10^6$  per ml with PBS without  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ ;
- iii) Add linearized DNA, 1-30  $\mu\text{g}$  per  $1\text{-}3 \times 10^5$  blastodermal cells in an eppendorf tube at room temperature. Add equimolar molar amounts of the non-linearized transgene plasmid bearing an attB site, and an integrase expression plasmid;
- 30 iv) Incubate at room temperature for 10 mins;

- v) Aliquot 100  $\mu$ l of the DNA-cell mixture to a 0.1 cm cuvette at room temperature;
- vi) Electroporate at 240 V and 25  $\mu$ FD (or 100 V and 125  $\mu$ FD for quail cells) using, for example, a Gene Pulser II<sup>TM</sup> (BIO-RAD).
- 5 vii) Incubate the cuvette at room temperature for 1-10 mins.
- viii) Before the electroporated cells are injected into a recipient embryo, they are transferred to a eppendorf tube at room temperature. The cuvette is washed with 350  $\mu$ l of media, which is transferred to the eppendorf, spun at room temperature and re-suspended in 0.01-0.3 ml medium;
- 10 ix) Inject 1-10  $\mu$ l of cell suspension into the subgerminal cavity of an non-irradiated or, preferably, an irradiated (e.g., with 300-900 rads) stage X egg. Shell and shell membrane are removed and, after injection, resealed according to U.S. Patent No. 6,397,777 incorporated herein by reference in its entirety; and
- 15 x) The egg is then incubated to hatching.

*(c) Blastodermal Cell Culture Medium:*

- i) 409.5 mls DMEM with high glucose, L-glutamine, sodium pyruvate, pyridoxine hydrochloride;
- ii) 5 mls Men non-essential amino acids solution, 10 mM;
- 20 iii) 5 mls Penicillin-streptomycin 5000 U/ml each;
- iv) 5 mls L-glutamine, 200 mM;
- v) 75 mls fetal bovine serum; and
- vi) 0.5 mls  $\beta$ -mercaptoethanol, 11.2mM.

25 **Example 5: Transfection of stage X embryos with attB plasmids**

*(a) DNA-PEI:* Twenty-five  $\mu$ g of a phage phiC31 integrase expression plasmid (pCMV-int), and 25  $\mu$ g of a luciferase-expressing plasmid (p $\beta$ -actin-GFP-attB) are combined in 200  $\mu$ l of 28 mM Hepes (pH 7.4). The DNA/Hepes is mixed with an equal volume of PEI which has been diluted 10-fold with water. The DNA/Hepes/PEI is incubated at room temperature for 15 mins. Three to seven  $\mu$ l of the complex are injected into the subgerminal cavity of windowed stage X white leghorn eggs which

are then sealed and incubated as described in U.S. Patents No. 6,397,777. The complexes will also be incubated with blastodermal cells isolated from stage X embryos which are subsequently injected into the subgerminal cavity of windowed irradiated stage X white leghorn eggs. Injected eggs are sealed and incubated as 5 described above.

(b) *Adenovirus-PEI*:

Two  $\mu$ g of a phage phiC31 integrase expression plasmid (pCMV-int), 2  $\mu$ g of a GFP expressing plasmid (p $\beta$ -actin-GFP-attB) and 2  $\mu$ g of a luciferase expressing plasmid (pGLB) were incubated with 1.2  $\mu$ l of JetPEI<sup>TM</sup> in 50  $\mu$ l of 20 mM Hepes buffer 10 (pH7.4). After 10 mins at 25°C, 3 x 10<sup>9</sup> adenovirus particles (Ad5-Null, Qbiogene) were added and the incubation continued for an additional 10 mins. Embryos are transfected *in ovo* or *ex ovo* as described above.

**Example 6: Stage I cytoplasmic injection**

15 Production of transgenic chickens by cytoplasmic DNA injection using DNA injection directly into the germinal disk as described in Sang *et al.*, *Mol. Reprod. Dev.*, 1: 98-106 (1989); Love *et al.*, *Biotechnology*, 12: 60-63 (1994) incorporated herein by reference in their entireties.

20 In the method of the present invention, fertilized ova, and preferably stage I embryos, are isolated from euthanized hens 45 mins. to 4 hrs. after oviposition of the previous egg. Alternatively, eggs were isolated from hens whose oviducts have been fistulated according to the techniques of Gilbert & Wood-Gush, *J. Reprod. Fertil.*, 5: 451-453 (1963) and Pancer *et al.*, *Br. Poult. Sci.*, 30: 953-7 (1989) incorporated herein in their entireties.

25 An isolated ovum was placed in dish with the germinal disk upwards. Ringer's buffer medium was then added to prevent drying of the ovum. Any suitable microinjection assembly and methods for microinjecting and reimplanting avian eggs are useful in the method of cytoplasmic injection of the present invention. A particularly suitable apparatus and method for use in the present invention is described 30 in U.S. Patent Application Serial No: 09/919,143 ("the '143 Application) and incorporated herein by reference in its entirety. The avian microinjection system

described in the '143 Application allowed the loading of a DNA solution into a micropipette, followed by prompt positioning of the germinal disk under the microscope and guided injection of the DNA solution into the germinal disk. Injected embryos could then be surgically transferred to a recipient hen as described, for 5 example, in Olsen & Neher, *J. Exp. Zool.*, 109: 355-66 (1948) and Tanaka *et al.*, *J. Reprod. Fertil.*, 100: 447-449 (1994). The embryo was allowed to proceed through the natural *in vivo* cycle of albumin deposition and hard-shell formation. The transgenic embryo is then laid as a hard-shell egg which was incubated until hatching 10 of the chick. Preferably, injected embryos were surgically transferred to recipient hens via the ovum transfer method of Christmann *et al.* in PCT/US01/26723, the contents of which are incorporated by reference in its entirety, and hard shell eggs were incubated and hatched.

Approximately 25 nl of DNA solution with either integrase mRNA or protein were injected into a germinal disc of stage I White Leghorn embryos obtained 90 15 minutes after oviposition of the preceding egg. Typically the concentration of integrase mRNA used was 100 ng/μl, and the concentration of integrase protein was 66 ng/μl.

To synthesize the integrase mRNA, a plasmid template encoding the integrase protein was linearized at the 3' end of the transcription unit. mRNA was synthesized, 20 capped and a polyadenine tract added using the mMESSAGE mMACHINE T7 Ultra Kit™ (Ambion, Austin, TX). The mRNA was purified by extraction with phenol and chloroform and precipitated with isopropanol. The integrase protein was expressed in *E. coli* and purified as described by Thorpe *et al.*, *Mol. Microbiol.*, 38: 232-241 (2000).

25 A plasmid encoding for the integrase protein is transfected into the target cells. However, since the early avian embryo transcriptionally silent until it reaches about 22,000 cells, injection of the integrase mRNA or protein was expected to result in better rates of transgenesis, as shown in the Table 1 below.

The chicks produced by this procedure were screened for the presence of the 30 injected transgene using a high throughput PCR-based screening procedure as described in Harvey *et al.*, *Nature Biotech.*, 20: 396-399 (2002).

Table 1: Summary of cytoplasmic injection results using different integrase strategies

Experimental group	Ovum transfers	Hard shells produced (%)	Chicks hatched (%) *	Transgenic chicks (%) ‡
No Integrase	5164	3634 (70%)	500 (14%)	58 (11.6%)
Integrase mRNA	1109	833 (75%)	115 (13.8%)	19 (16.5%)
Integrase protein	374	264 (70.6%)	47 (17.8%)	16 (34%)

\* : Percentages based on the number of hard shells

‡ : Percentages based on the number of hatched birds

5

**Example 7: Characterization of phiC31 integrase-mediated integration sites in the chicken genome.**

To characterize phiC31-mediated integration into the chicken genome, a plasmid rescue method was used to isolate integrated plasmids from transfected and 10 selected chicken fibroblasts. Plasmid pCR-XL-TOPO-CMV-pur-attB (SEQ ID NO: 10, shown in Fig. 18) does not have *Bam*H I or *Bgl* II restriction sites. Genomic DNA from cells transformed with pCR-XL-TOPO-CMV-pur-attB was cut with *Bam*H I or *Bgl* II (either or both of which would cut in the flanking genomic regions) and 15 religated so that the genomic DNA surrounding the integrated plasmid would be captured into the circularized plasmid. The flanking DNA of a number of plasmids were then sequenced.

DF-1 cells (chicken fibroblasts),  $4 \times 10^5$  were transfected with 50 ng of pCR-XL-TOPO-CMV-pur-attB and 1  $\mu$ g of pCMV-int. The following day, the culture medium was replaced with fresh media supplemented with 1  $\mu$ g/ml puromycin. After 20 10 days of selection, several hundred puromycin-resistant colonies were evident. These were harvested by trypsinization, pooled, replated on 10 cm plates and grown to confluence. DNA was then extracted.

Isolated DNA was digested with *Bam*H I and *Bgl* II for 2-3 hrs, extracted with phenol:chloroform:isoamyl alcohol chloroform:isoamyl alcohol and ethanol

precipitated. T4 DNA ligase was added and the reaction incubated for 1 hr at room temperature, extracted with phenol:chloroform:isoamyl alcohol and chloroform:isoamyl alcohol, and precipitated with ethanol. 5  $\mu$ l of the DNA suspended in 10 $\mu$ l of water was electroporated into 25  $\mu$ l of Genehogs<sup>TM</sup> (Invitrogen) 5 in an 0.1 cm cuvette using a GenePulser II (Biorad) set at 1.6 kV, 100 ohms, 25 uF and plated on Luria Broth (LB) plates with 5  $\mu$ g/ml phleomycin (or 25  $\mu$ g/ml zeocin) and 20  $\mu$ g/ml kanamycin. Approximately 100 individual colonies were cultured, the plasmids extracted by standard miniprep techniques and digested with *Xba* I to identify clones with unique restriction fragments.

10       Thirty two plasmids were sequenced with the primer attB-for (5'-TACCGTCGACGATGTAGGTACGGTC-3') (SEQ ID NO: 12) which allows sequencing across the crossover site of attB and into the flanking genomic sequence. All of plasmids sequenced had novel sequences inserted into the crossover site of attB, indicating that the clones were derived from plasmid that had integrated into the 15 chicken genome via phiC31 integrase-mediated recombination.

The sequences were compared with sequences at GenBank using Basic Local Alignment Search Tool (BLAST). Most of the clones harbored sequences homologous to *Gallus* genomic sequences in the TRACE database.

20       Example 8: Insertion of a wild-type attP site into the avian genome augments integrase-mediated integration and transgenesis.

The chicken B-cell line DT40 cells (Buerstedde *et al.*, *E.M.B.O. J.*, 9: 921-927 (1990)) are useful for studying DNA integration and recombination processes (Buerstedde & Takeda, *Cell*, 67:179-88 (1991)). DT40 cells were engineered to 25 harbor a wild-type attP site isolated from the *Streptomyces* phage phiC31. Two independent cell lines were created by transfection of a linearized plasmid bearing an attP site linked to a CMV promoter driving the resistance gene to G418 (DT40-NLB-attP) or bearing an attP site linked to a CMV promoter driving the resistance gene for puromycin (DT40-pur-attP). The transfected cells were cultured in the presence of 30 G418 or puromycin to enrich for cells bearing an attP sequence stably integrated into the genome.

A super-coiled luciferase vector bearing an attB (SEQ ID NO: 2 shown in Fig. 10) was co-transfected, together with an integrase expression vector CMV-C31int (SEQ ID NO: 1) or a control, non-integrase expressing vector (CMV-BL) into wild-type DT40 cells and the stably transformed lines DT40-NLB-attP and DT40-pur-attP.

5 Cells were passaged at 5, 7 and 14 days post-transfection and about one third of the cells were harvested and assayed for luciferase. The expression of luciferase was plotted as a percentage of the expression measured 5 days after transfection. As can be seen in Fig. 21, in the absence of integrase, or in the presence of integrase but in the DT40 cells lacking an inserted wild-type attP site, luciferase expression from a 10 vector bearing attB progressively decreased to very low levels. However, luciferase levels were persistent when the luciferase vector bearing attB was co-transfected with the integrase expression vector into the attP bearing cell lines DT40-NLB-attP and DT40-pur-attP. Inclusion of an attP sequence in the avian genome augments the level 15 of integration efficiency beyond that afforded by the utilization of endogenous pseudo-attP sites.

**Example 9: Generation of attP transgenic cell line and birds using an NLB vector**

The NLB-attP retroviral vector can be injected into stage X chicken embryos 20 laid by pathogen-free hens. A small hole is drilled into the egg shell of a freshly laid egg, the shell membrane cut away and the embryo visualized by eye. With a drawn needle attached to a syringe, 1 to 10  $\mu$ l of concentrated retrovirus, approximately  $2.5 \times 10^5$  IU, is injected into the subgerminal cavity of the embryo. The egg shell is resealed with a hot glue gun. Suitable methods for the manipulation of avian eggs, including 25 opening and resealing hard shell eggs are described in U.S. Patent Serial Nos: 5,897,998 and 6,397,777 which are herein incorporated by reference in their entireties.

Typically, 25% of embryos hatch 21 days later. The chicks are raised to sexual 30 maturity and semen samples are taken. Birds that have a significant level of the transgene in sperm DNA will be identified, typically by a PCR-based assay. Ten to 25% of the hatched roosters will be able to give rise to G1 transgenic offspring, 1 to 20% of which may be transgenic. DNA extracted from the blood of G1 offspring is

analyzed by PCR and Southern analysis to confirm the presence of the intact transgene. Several lines of transgenic roosters, each with a unique site of attP integration, are then bred to non-transgenic hens, giving 50% of G2 transgenic offspring. Transgenic G2 hens and roosters from the same line can be bred to produce 5 G3 offspring homozygous for the transgene. Homozygous offspring will be distinguished from hemizygous offspring by quantitative PCR. The same procedure can be used to integrate an attB or attP site into transgenic birds.

**Example 10: Expression of immunoglobulin chain polypeptides by transgenic 10 chickens**

Bacterial artificial chromosomes (BACs) containing a 70 kbp segment of the chicken ovomucoid gene with the light and heavy chain cDNAs for a human monoclonal antibody inserted along with an internal ribosome entry site into the 3' untranslated region of the ovomucoid gene were equipped with the attB sequence. 15 The heavy and light chain cDNAs were inserted into separate ovomucoid BACs such that expression of an intact monoclonal antibody requires the presence of both BACs in the nucleus.

Several hens produced by coinjection of the attB-bearing ovomucoid BACs and integrase-encoding mRNA into stage I embryos produced intact monoclonal 20 antibodies in their egg white. One hen, which had a high level of the light chain ovomucoid BAC in her blood DNA as determined by quantitative PCR particularly expressed the light chain portion of the monoclonal antibody in the egg white at a concentration of 350 nanograms per ml, or approximately 12 µg per egg.

**What is Claimed Is:**

1. A method for genetically transforming an avian cell, comprising:
  - 5 delivering to an avian cell having a first recombination site a nucleic acid molecule comprising a second recombination site;
  - delivering a source of integrase activity to the avian cell; and
  - 10 maintaining the avian cell under conditions suitable for the integrase to mediate recombination between the first and the second recombination sites, thereby integrating the nucleic acid molecule into the nuclear genome of the avian cell.
- 15 2. The method of Claim 1, wherein the avian cell is selected from the group consisting of a stage I cell, a stage X blastodermal cell, a primordial germline cell and an oviduct cell.
3. The method of Claim 1, wherein the integrase is a serine recombinase
4. The method of Claim 1, wherein the integrase is the bacteriophage phiC31 integrase.
- 20 5. The method of Claim 1, wherein the source of integrase activity is delivered to the avian cell as a polypeptide.
6. The method of Claim 1, wherein the source of integrase activity is expressed from a polynucleotide molecule delivered to the avian cell.
- 25 7. The method of Claim 6, wherein the polynucleotide molecule is an mRNA.
8. The method of Claim 6, wherein the polynucleotide molecule is an expression vector.

30

9. The method of Claim 8, wherein the expression vector comprises a promoter operably linked to a region encoding the integrase, wherein the promoter is selected from a CMV promoter and an RSV promoter.
- 5 10. The method of Claim 9, wherein the expression vector comprises a nucleotide sequence selected from SEQ ID NOs: 1 and 9.
11. The method of Claim 1, wherein the nucleic acid molecule comprises an isolated gene expression control region operably linked to a region encoding a 10 polypeptide selected to be expressed by an avian cell.
12. The method of Claim 1, wherein the nucleic acid molecule is an expression vector.
- 15 13. The method of Claim 12, wherein the expression vector is selected from the group consisting of a plasmid, an artificial chromosome, an isolated eukaryotic chromosome, an adenovirus vector and a lentivirus vector.
- 20 14. The method of Claim 11, wherein the gene expression control region is isolated from an avian gene selected from the group consisting of an ovalbumin gene, a lysozyme gene and an ovomucoid gene.
- 25 15. The method of Claim 11, wherein the nucleic acid molecule further comprises an IRES operably linked to a region encoding a second polypeptide selected to be expressed by an avian cell.
16. The method of Claim 11, wherein the region encoding the polypeptide is codon-optimized for expression of the polypeptide by the avian cell.
- 30 17. The method of Claim 15 wherein the region encoding the second polypeptide is codon-optimized for expression of the polypeptide by the avian cell.

18. The method of Claim 1, wherein the first recombination site is a attP site having at least 25% identity to the sequence according to SEQ ID NO: 11.
- 5 19. The method of Claim 1, wherein the first recombination site is a attP site having at least 75% identity to the sequence according to SEQ ID NO: 11.
20. The method of Claim 1, wherein the first recombination site is a attP site according to SEQ ID NO: 11.
- 10 21. The method of Claim 1, wherein the first recombination site is a pseudo-attP site.
- 15 22. The method of Claim 1, wherein the nucleic acid molecule and the source of integrase activity are sequentially delivered to the isolated avian cell.
23. The method of Claim 1, wherein the nucleic acid molecule and the source of integrase activity are concurrently delivered to the isolated avian cell.
- 20 24. The method of Claim 1, wherein the avian cell is a chicken cell.
25. The method of Claim 1, wherein the avian cell is a quail cell.
26. The method of Claim 1, wherein the nucleic acid molecule and the source of integrase activity are delivered to the avian cell by a method or combination of methods selected from injection, electroporation, DNA-PEI, virus-PEI, microcell fusion, artificial viral coats, polylysine conjugation and lipofection.
- 30 27. The method of Claim 1, wherein the avian cell is an oviduct cell in the oviduct of an avian and further comprising the steps of:

before delivering the nucleic acid molecule and source of integrase activity, surgically exposing the luminal surface of the oviduct of an avian;

5                   delivering the nucleic acid molecule and the source of integrase activity to the oviduct cell by depositing the nucleic acid molecule and the source of integrase activity on the luminal surface;

10                  applying an electroporation pulse across the wall of the oviduct, wherein the electroporation pulse is directed to deliver the nucleic acid molecule and the source of integrase activity to recipient oviduct cells, whereby the integrase mediates integration of the nucleic acid molecule into the genome of a recipient cell;

15                  surgically closing the oviduct; and

                        maintaining the avian under conditions suitable for the expression by the recipient oviduct cells of a polypeptide encoded by the nucleic acid molecule and deposition of said polypeptide into the white of a laid hard shell egg.

28.                A genetically transformed avian cell, and progeny thereof, generated by a of the methods according to Claims 1.

20

29.                A genetically transformed avian, and progeny thereof, generated by the method according to Claim 27.

30.                A method for generating a genetically transformed avian, comprising:

25                  genetically transforming an avian cell by a method according to Claim 1;

                        delivering the genetically transformed avian cell to an early stage avian embryo, thereby generating a genetically transformed embryo; and

30                  maintaining the avian embryo under conditions suitable for the embryo to develop and hatch as a transgenic chick.

31. The method of Claim 26, wherein the genetically transformed embryo is delivered to a recipient female avian and laid as a hard shell egg.
- 5 32. A method of producing a polypeptide comprising expressing the polypeptide in a transgenic avian cell, said cell comprising a heterologous nucleic acid molecule that expresses the polypeptide, wherein the heterologous nucleic acid molecule is integrated into the nuclear genome and flanked by halves of an integration site recognized by an integrase.
- 10 33. A polypeptide of interest expressed by a genetically transformed avian cell, and progeny thereof, the cell comprising a heterologous nucleic acid molecule integrated into the nuclear genome thereof by an integrase, and wherein the nucleic acid molecule comprises a region encoding the polypeptide of interest and an operably linked promoter region.
- 15 34. The polypeptide according to Claim 33, wherein the polypeptide is an immunoglobulin chain.
- 20 35. The polypeptide according to Claim 33, wherein the polypeptide is a cytokine.
36. The polypeptide according to Claim 33, wherein the polypeptide is deposited in the white of an avian egg.
- 25 37. The polypeptide according to Claim 33, wherein the polypeptide is expressed by a transgenic bird comprising the genetically transformed avian cell.
38. The polypeptide according to Claim 37, wherein the genetically transformed avian cell, and the progeny thereof, are generated according to the methods of Claim 1.

39. The polypeptide according to Claim 37, wherein the transgenic bird is generated according to the method of Claim 30.
40. An avian egg white comprising a heterologous polypeptide produced according to the method according to the Claim 32.
41. An avian egg comprising a heterologous polypeptide produced according to the method according to the Claim 32.
- 10 42. A method for generating a genetically modified avian cell, and progeny thereof, using a tagged chromosome, said method comprising the steps of:
  - providing an isolated modified chromosome comprising a lac operator region and a first recombination site;
  - delivering the modified chromosome to a avian cell, thereby generating a trisomic avian cell;
  - delivering to the avian cell a source of a tagged polypeptide comprising a fluorescent domain and a lac repressor domain;
  - delivering a source of integrase activity to the avian cell;
  - delivering a polynucleotide comprising a second recombination site and a region encoding a polypeptide to the avian cell;
  - 20 maintaining the avian cell under conditions suitable for the integrase to mediate recombination between the first and second recombination sites, thereby integrating the polynucleotide into the modified chromosome and generating a genetically modified avian cell;
  - 25 expressing the tag polypeptide by the avian cell;
  - allowing the tag polypeptide to bind to the modified chromosome so as to label the modified chromosome; and
  - isolating the modified chromosome by selecting modified chromosomes having a tag polypeptide bound thereto.

43. The method according to Claim 42, further comprising the step of delivering the modified chromosome to a second avian cell to generate a trisomic avian cell.
44. A modified avian chromosome comprising a lac operator region and a recombination site.  
5
45. The avian chromosome according to Claim 44, wherein the lac operator region is a concatamer of lac operators.
46. The avian chromosome according to Claim 44, wherein the recombination site is selected from an att B or an att P site.
- 10 47. The avian chromosome according to Claim 44, further comprising a region encoding a tag polypeptide and an operably linked promoter.
48. The avian chromosome according to Claim 47, wherein the tag polypeptide comprises a fluorescent domain and a lac repressor domain.
- 15 49. The avian chromosome according to Claim 44, wherein the fluorescent domain is GFP.
50. The avian chromosome according to Claim 45, further comprising a heterologous polynucleotide integrated into the first recombination site.
- 20 51. The avian chromosome according to Claim 50, wherein the heterologous polynucleotide comprises a region encoding a polypeptide operably linked to a promoter.
52. The avian chromosome according to Claim 50, wherein the promoter is an avian promoter selected from the gene expression control region of an onomucoid gene, a lysozyme gene and an ovalbumin gene.

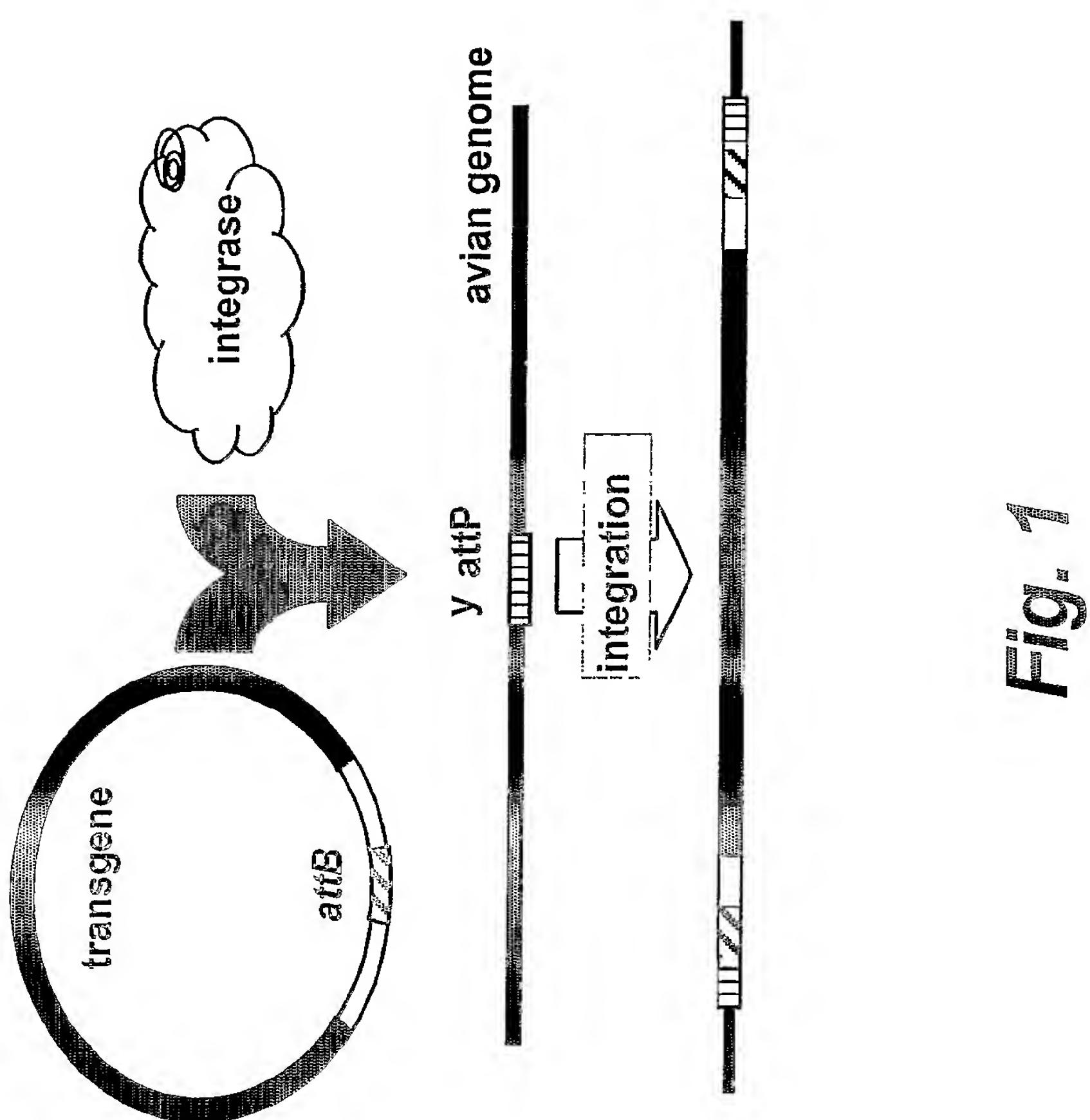


Fig. 1

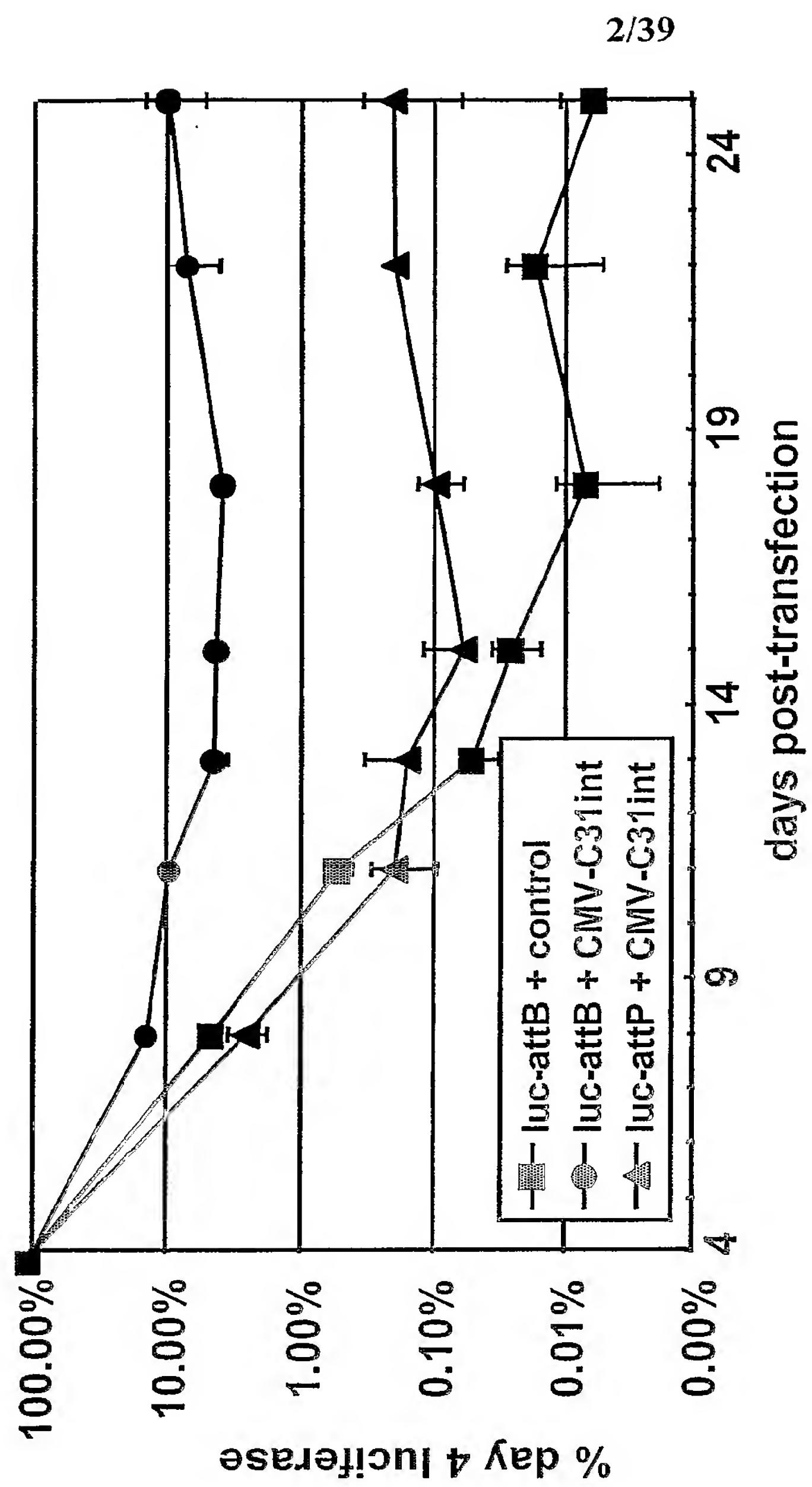
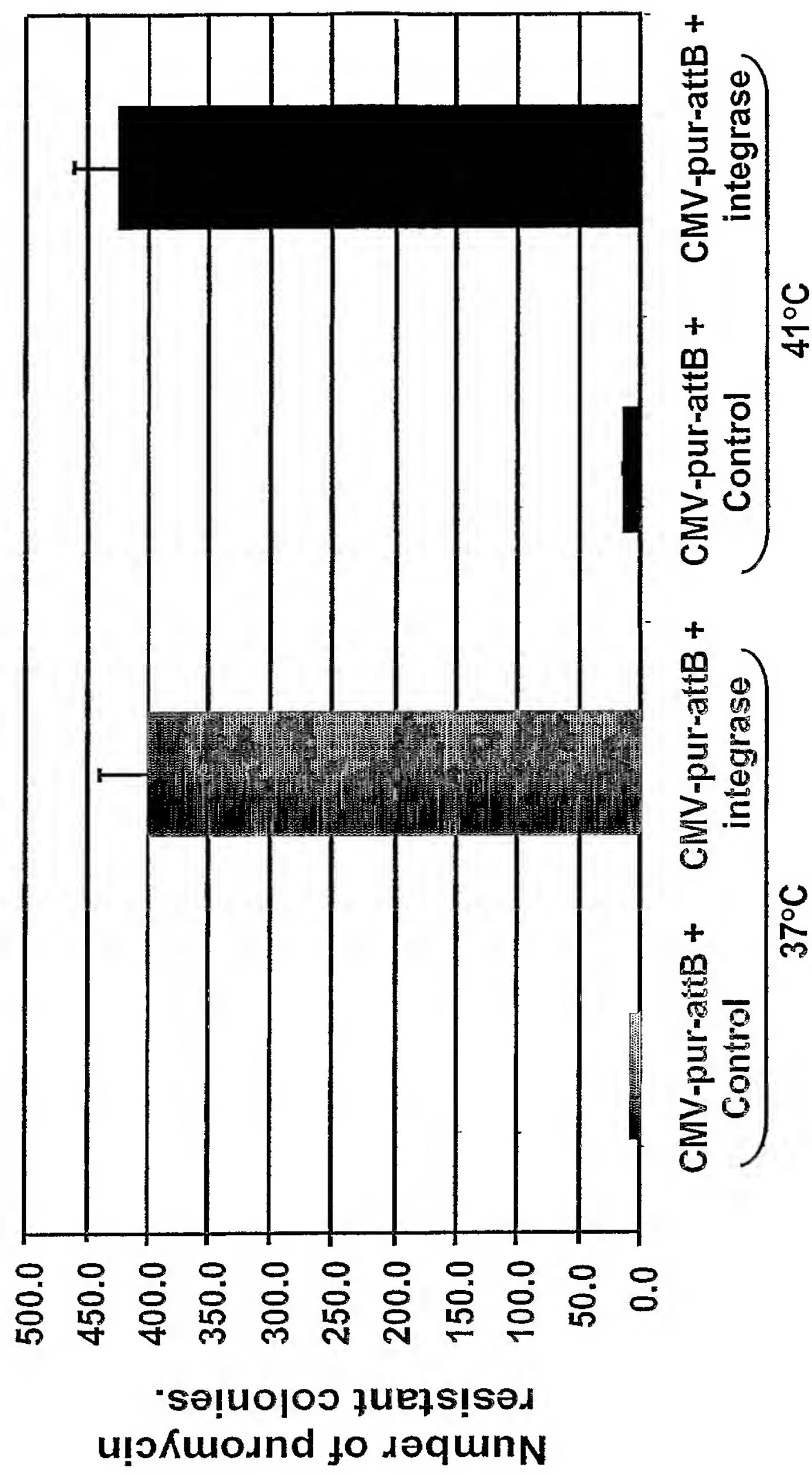


Fig. 2

3/39



### Fig. 3

4/39

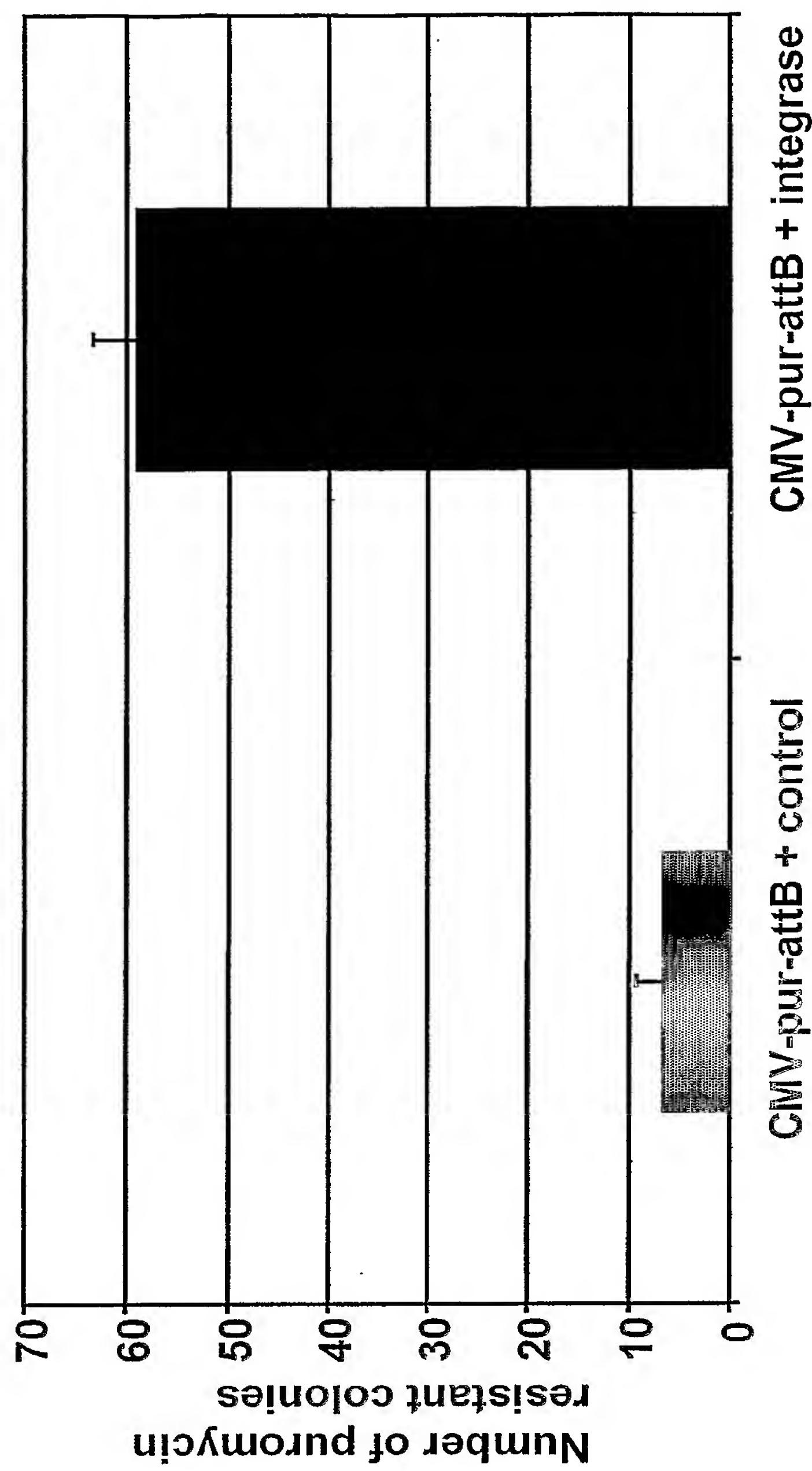
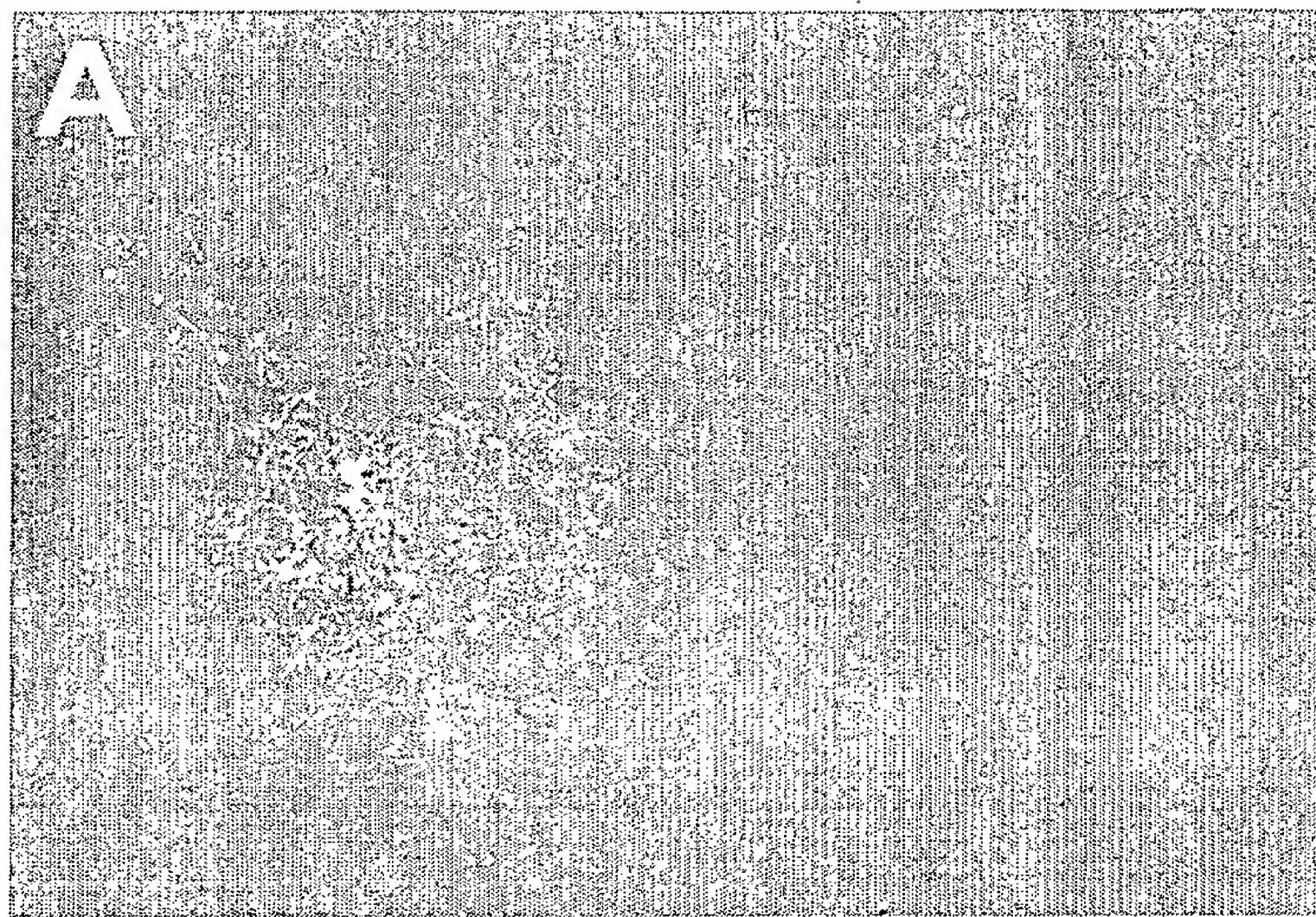


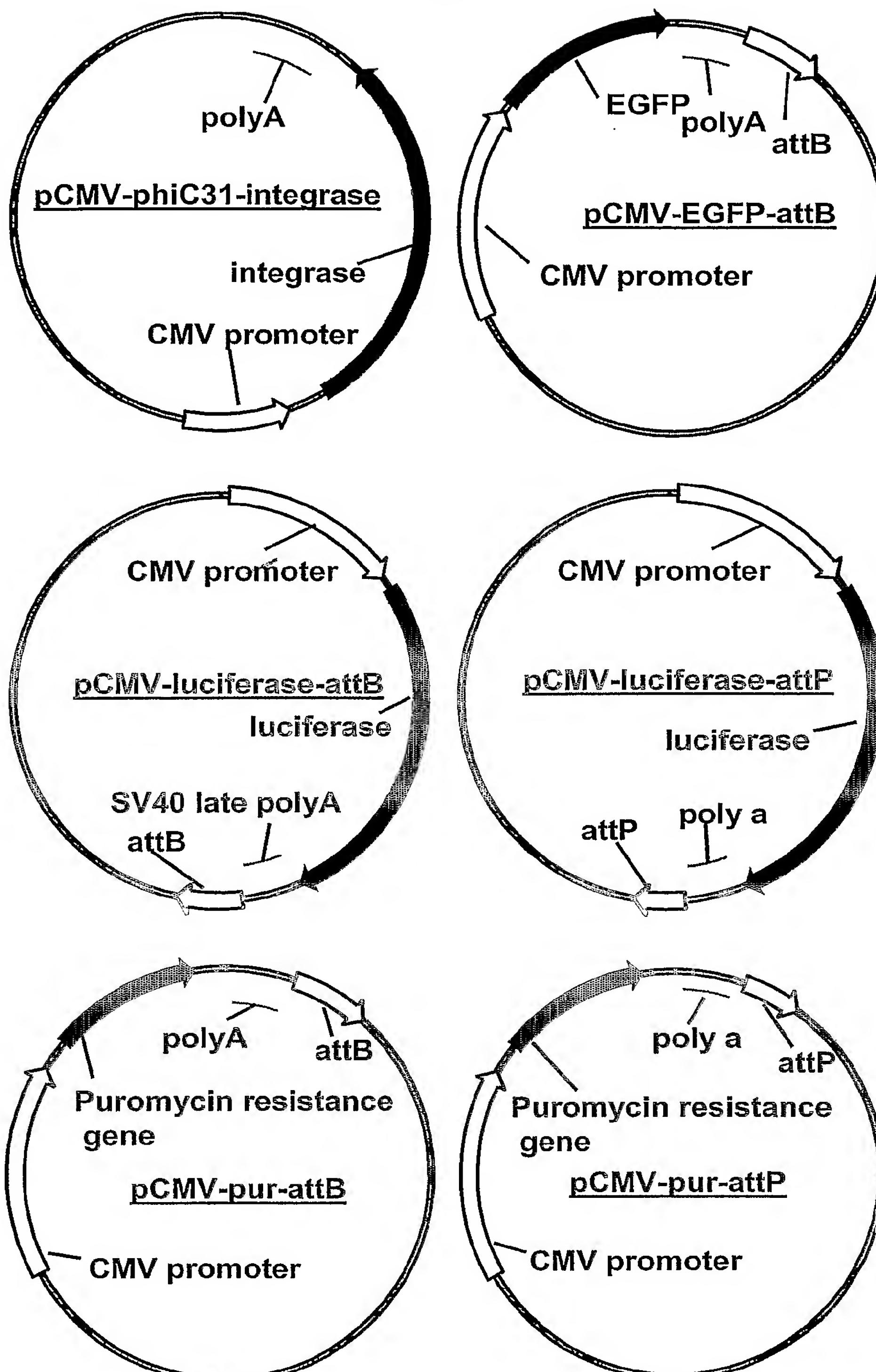
FIG. 4

5/39

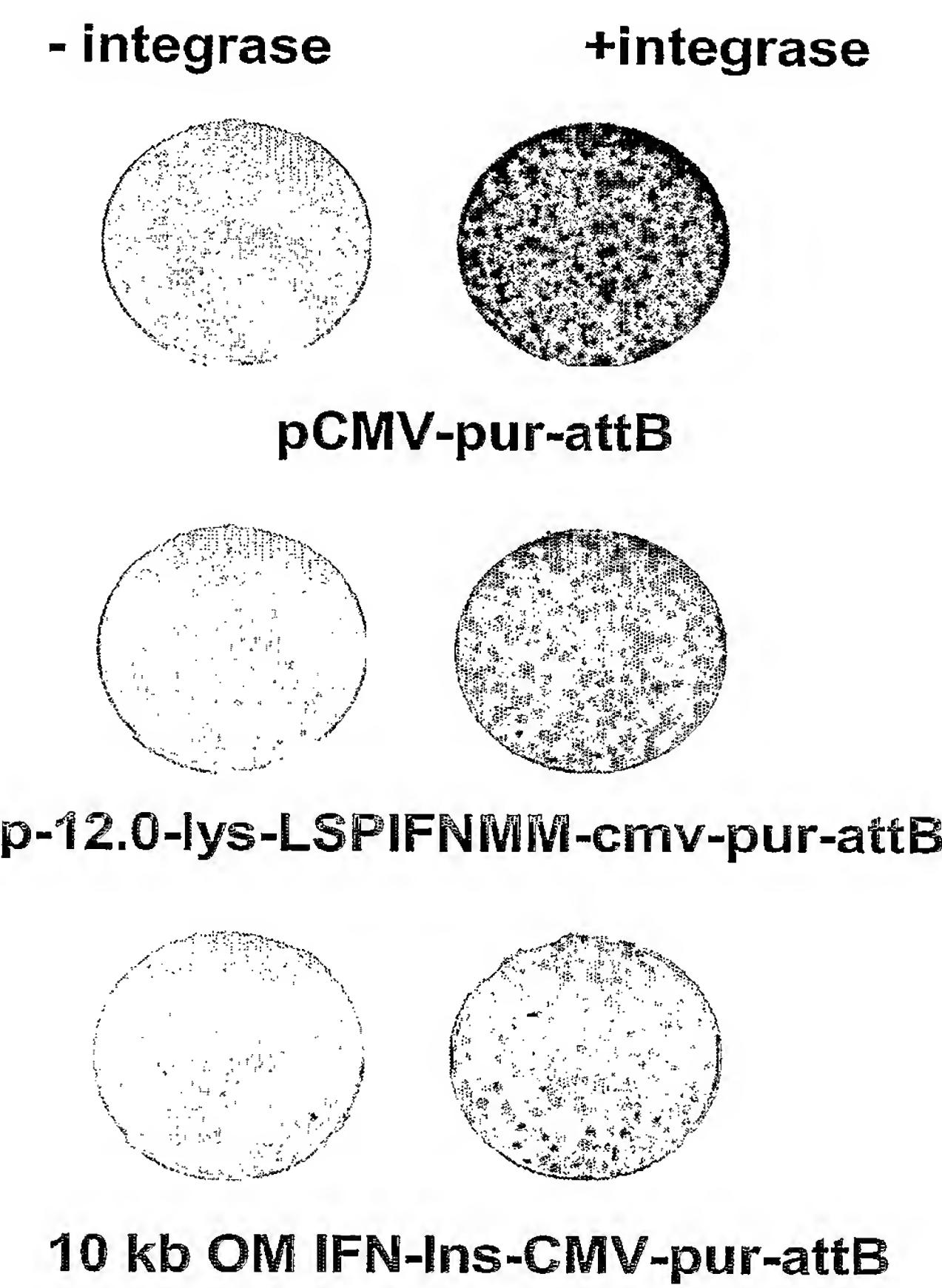


*Fig. 5*

6/39

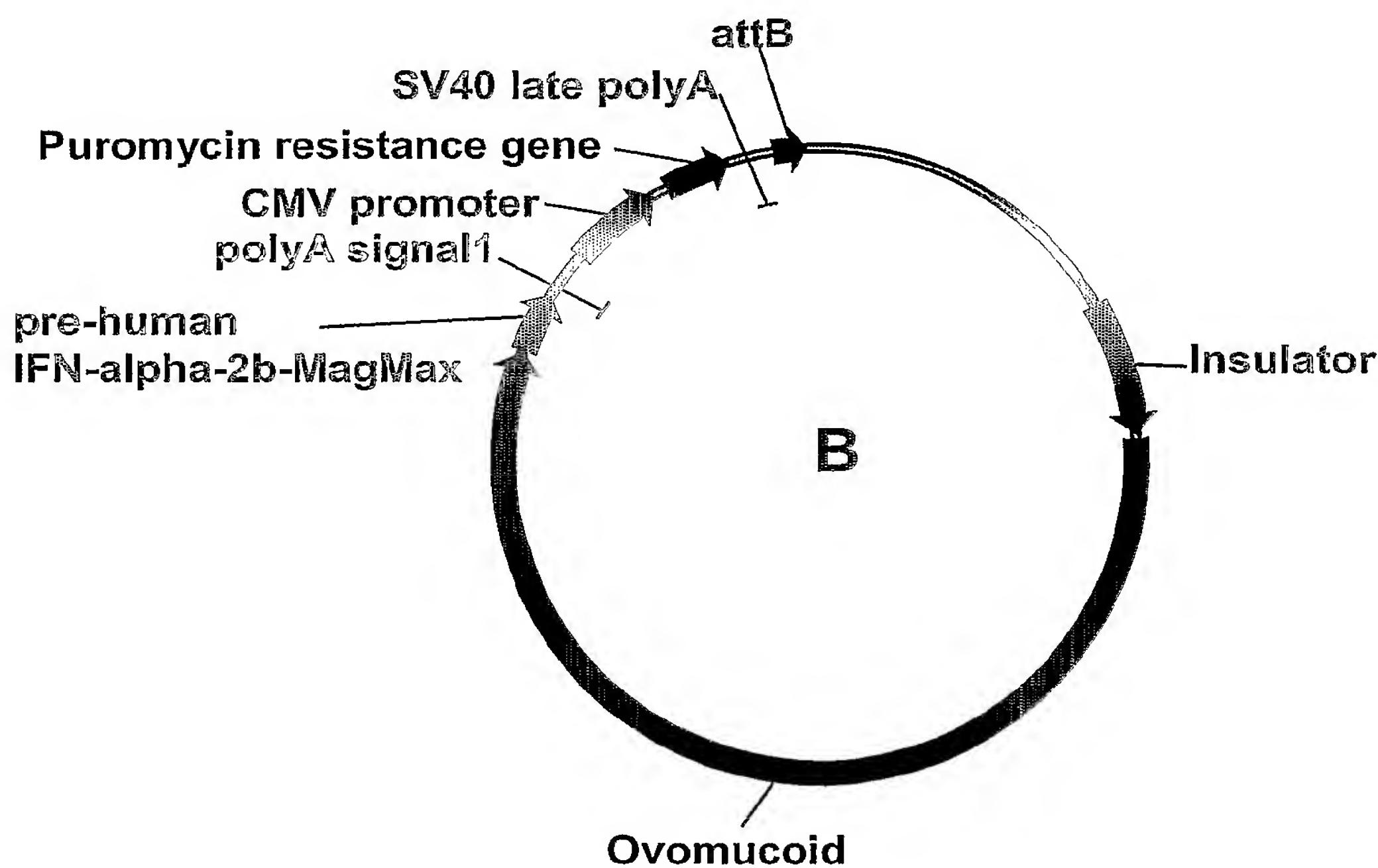
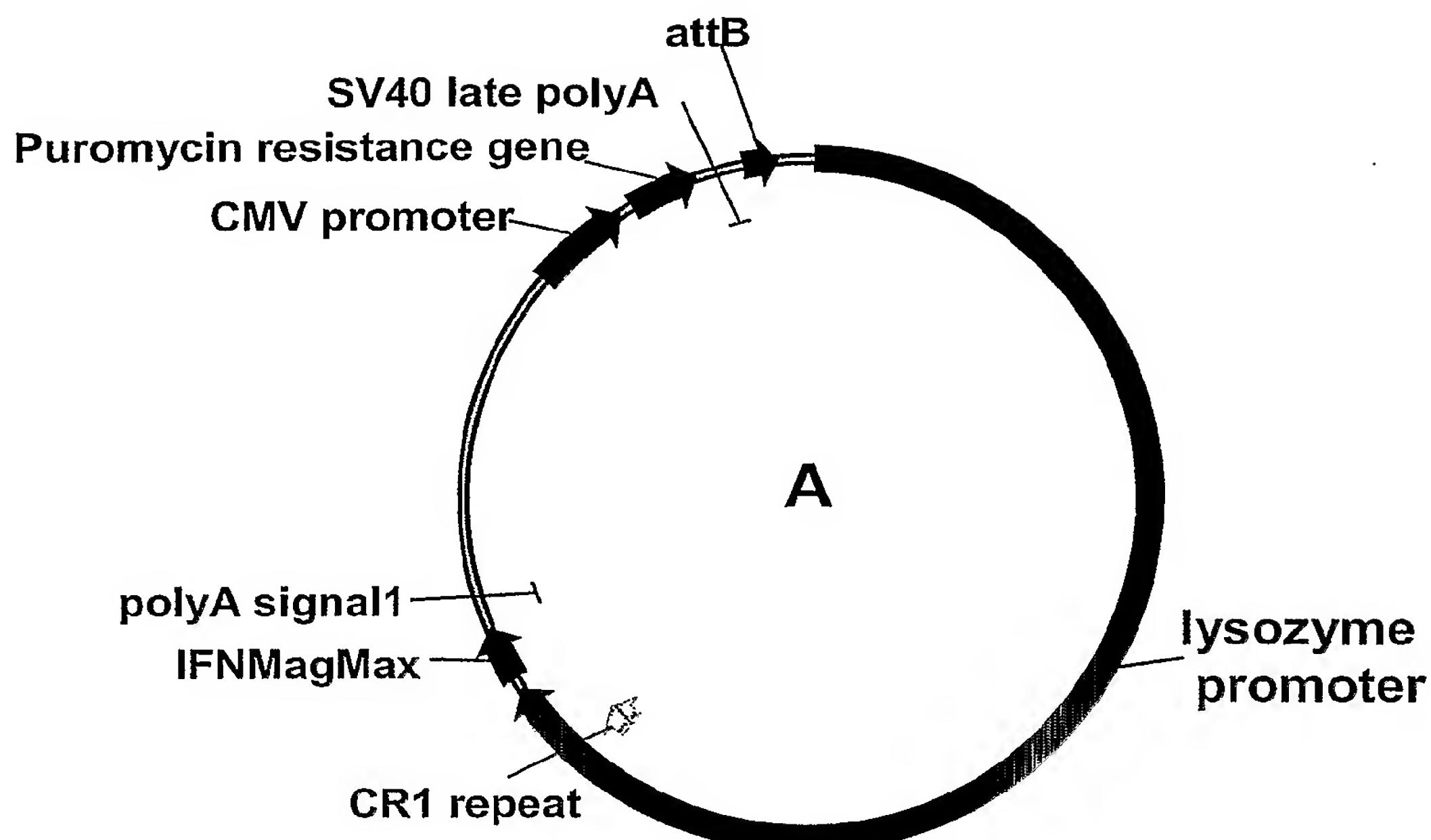
*Fig. 6*

7/39



***Fig. 7***

8/39



**Fig. 8**

9/39

**pCMV-C31int (SEQ ID NO: 1)**

CATTCGCCATTCAAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGCCTTCGCTATT  
ACGCCAGCCAATACGCAAACCGCCTCTCCCCCGCGTTGGCCGATTCAATTAGCAGGATCG  
ATCCAGACATGATAAGATACTTGATGAGTTGGACAAACACAAACTAGAATGCAGTGAAAA  
AAATGCTTATTGTGAAATTGTGATGCTATTGCTTATTGTAAACCATTATAAGCTGCAA  
TAAACAAGTTAACAAACAATTGCATTCAATTATGTTCAGGTTAGGGGAGGTGTGGG  
AGGTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGATCATGAACAG  
ACTGTGAGGACTGAGGGCCTGAAATGAGCCTGGGACTGTGAATCTAAAATACACAAACAA  
TTAGAATCACTAGCTCCTGTGTATAATATTTCATAAAATCATACTCAGTAAGCAAAACCTC  
AAGCAGCAAGCATATGCAGCTAGTTAACACATTACACTTAAAAATTATATTACCTT  
AGAGCTTAAATCTCTGTAGGTAGTTGCTAACATTATGTACACACCAGAAGTAAGGTCCT  
TCACAAAGATCCCAGCTAGCTATAATACGACTCACTATAGGGAGAGAGCTATGACGTCGC  
ATGCACGCGTAAGCTTGGGCCCTCGAGGGATCCGGGTGTCTCGCTACGCCGCTACGTCTC  
CGTGCCTCTCGATGGCGTTCCCTGCCCCCTGCCCGTAGTCGACTTCGTGACAACGATCTG  
TCTACGAAGAGCCGACGAACACCGCCTGTCGTACTGACGCGCCCGACCACGACTT  
AGGGCCGGTCGGGTCAAGCGTCGGCGTCTCGGGGAACCATTGGTCAAGGGGAAGCTCGGGG  
CTTCGGCGGCTTCAAGTCGGCAAGCCGCTCTCCGCCCCCTGCTGCCGGAGCGTCAGCGCT  
GCCTGTTGCTTCCGGAAGTGCTTCCTGCCAACGGGTCCGTCGTACGCGCCTGCCGCGGTC  
TTCGTACAGCTCTCAAGGGCGTTCAAGGGCGTCCAGGGCGTCGGCGCCTCCGCAACAAAGG  
CGCCGCTCTCTCAGGCGCTCAGTGAGCTGCCAAGCGTCGGGCTTCCCACAGAAC  
GCCAACGTCTTCGTCGCCTCGCGTGCCTGATCTTGTGAAGATGCGTTCCGCAACGAA  
CTTGTGAGTGCCGCATGCTGACGTTGCACGTGCCCTCGTGTGCCAGGTGCGGACGGGT  
CGACCACCTCCGGCGACGGCAGCGTAAGAGTCCTGATCGATTCTCCCCGCGCTTCGAA  
GTCATGACGGGCCACACTCGCAGTACAGCTGTCCATGGCGGACAGAATGGCTGCCCG  
GGAAAGCCCTTGGCGGCCCTGCCGTCCAACCACGCCGAAGCTCATACCACTCAGCGG  
GCTCGATGATCGGTCCGCAATCAAGCTCGACCGGCCGGAGCGTGTGGTCGCGCTGAATG  
CGGTAACCCTCAATCTCGTGGTGGCGTCCGGCTCTTGTAGATCACCTCAGC  
GGCGAAGCCCGCAATACGCGGGCCCAGGGATTCCGATAACGGTTGCCGGTCCAGGCG  
TTGAAGCGGTCTTCTCCAATCGTCTGCCCGGGTGGCACGGCGTCAGCGTCCATGCGC  
TTACAAAGCCCCGTGATGCTGCCGGGTGAATGGCGGCTTGACTGCCGGCTTGAAGGG  
GTGTTGTGCGTCTTGATCTCACGCCACCACCGGATTACGTGGCTCGAACACTCGAAGG  
GTCCGGTAAGGGAGTGGTCGAGTGCACGCTTGTGATGACGACATTGACCATTGGCC  
TTGCGCGTGAATCTCCTCGTCTCCGAAACAAGCTCGAAGCCGTAAGGCCCTCCGCC  
GTACCCGCCAATCGCGCTGAAGGTTCTCGTGTGAGAATCTCGCCACTTCAGCGAAG  
ATTCTTGTGCGACCGTCAAGCGCATAATCAGGTGAATCAGGTCCATGACGTTCC  
CGGAAGACGCCCTCGTGAAGTGGAAACAATCGTCACGCCAGGGCGAGCAATTCC  
CGGAATCGCGTCCATGACCTTCAGGCGCGAGAAGCGCGACACGTCAAGACATGAT  
TGAGCCGCCGGCGCATTGTTCAAGGATGCGTTCAACTCCGGCGCTCCGCCGTCC  
AACGCCGACGTGCCCGCGCTCGTGAATGCCGACGAACCTGAACCGGCCCGTC  
CTCGACTTCGCGCTGAAGGTCGGCCGCCCTGCTTCGTTGGCGCTACGCTGTGCG  
TTGCTGCGCTCGAATTCTCGCGCTCGCGACTGACGGTCGTAAGCACCCCGTACGT  
ACCCCGGTACAACCCCTTGTGATGTCATGTCGGCGACCCCTACGACTAGTGAG  
GAGCTCGGACCGGTACCTGCAGGCGTACCTGACGGTTCACTAAACGAGCT  
GCAGTGGATGGCGTCTCCAGGCGATCTGACGGTTCACTAAACGAGCT  
CCCACCGTACACGCCCTACGCCATTGCGTCAATGGGGCGGAGTTGTTACGAC  
AAGTCCCCTGATTTGGTGCACAAACAAACTCCCATTGACGTCAATGGGGTGGAGACT  
AAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCAAGTAG  
GGAAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGT  
CCCATAAGGTCACTCCACCCATTGAC  
CATATGATACACTTGATGACTGCCAAGTGGCAGTTACCGTAAATACT  
CCACCCATTGAC

10/39

GTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGG  
 CGGGGGTCGTTGGCGGTAGCCAGGCGGCCATTACCGTAAGTTATGTAACGACCTGCAC  
 GATGCTGTTCTGTGAAATTGTTATCCGCTCACATTCCACACATTATACGAGCCGGAA  
 GCTATAAAAGTGTAAAGCCTGGGTGCCATTAGAGTGAAGGGCCTCGTATACGCTATT  
 ATAGGTTAATGTCATGATAATAATGGTTCTTAGACGTCAGGTGGCACTTTGGGGAAATG  
 TCGCGGAACCCCTATTGTTATTCTAAATACATTCAAATATGTATCCGCTCATGAGA  
 CAATAACCCTGATAAAATGCTCAATAATATTGAAAAACGCGCAATTGCAAGCTGCA  
 ATGAATCGGCCAACCGCGGGGAGAGGCAGGTTGCGTATTGGCGCTTCCGCTTCGC  
 TCACTGACTCGCTCGCTCGGTGTTGGCTGGCGAGCGGTATCAGCTCACTCAAAGGCG  
 GTAATACGGTTATCCACAGAATCAGGGATAACGCAAGGAAGAACATGTGAGCAAAAGGCCA  
 GCAAAAGGCCAGGAACCGTAAAAAGGCCGTTGCTGGCTTTCCATAGGCTCCGCC  
 CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA  
 AGATAACCAGGCAGGTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCGCT  
 TACCGGATACTGTCGCCCTTCTCCCTCGGAAGCGTGGCGCTTCTCAATGCTCACGCT  
 GTAGGTATCTCAGTTGGTGTAGGTCGCTCCAAGCTGGCTGTGACGAACCC  
 GTTCAGCCGACCGCTGCGCCTATCCGGTAACTATCGTCTTGAGTCCAACCCGTAAGACA  
 CGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG  
 GTGCTACAGAGTTGAGTGGCTGCTAACACTACGGCTACACTAGAAGGACAGTATTGGT  
 ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTTGTGATCCGGCAA  
 ACAAAACCACCGCTGGTAGCGGTGGTTTTGCAAGCAGCAGATTACGCGCAGAAAAAA  
 AAGGATCTCAAGAAGATCCTTGATCTTCTACGGGGTCTGACGCTCAGTGGAACGAAAAC  
 TCACGTTAAGGGATTTGGTCATGCCATACTTCGTATAGCATACTTACATTACGAAGTTATGG  
 CATGAGATTATCAAAAGGATCTCACCTAGATCCTTTAAATTAAAAATGAAGTTAAAT  
 CAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA  
 CCTATCTCAGCGATCTGCTATTGCTCATCCATAGTGCCTGACTCCCCGTGTTAGAT  
 AACTACGATAACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAACCGCAGACCCAC  
 GCTCACCGGCTCCAGATTATCAGCAATAAACAGCCAGCCGGAAAGGGCCAGCGCAGAAGT  
 GGTCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGTTGCCGGAAAGCTAGAGTAAG  
 TAGTTGCGCAGTTAATAGTTGCGCAACGTTGTCATTGCTACAGGCATCGTGGTGTAC  
 GCTCGTCGTTGGTATGGCTTCATTAGCTCCGGTCCACGATCAAGCGAGTTACATGA  
 TCCCCCATGTTGCAAAAAAGCGGTTAGCTCCTCGGTCCGATCGTGTAGAAGTAA  
 GTTGGCCGAGTGTATCACTCATGGTTATGGCAGCAGTCATAATTCTCTACTGTCT  
 CATCCGTAAGATGCTTTCTGTGACTGGTAGTACTCAACCAAGTCATTGAGAATAGTGT  
 ATGCGGCACCGAGTTGCTCTGCCCAGTCATAACGGGATAATTACCGGCCACATAGCAG  
 AACTTTAAAAGTGCCTCATCATTGGAAAACGTTCTCAGGGCGAAAACCTCAAGGATCTTAC  
 CGCTGTTGAGATCCAGTTGATGTAACCCACTCGTGCACCCACTGATCTCAGCATCTT  
 ACTTTCAACCAGCGTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGAAAAAGGGAAT  
 AAGGGCGACACGGAAATGTTGAATACTCATACTCTTCTTTCAATATTATTGAAGCATT  
 ATCAGGGTTATTGTCATGCCAGGGTGGCACACATATTGATAACCAGCGATCCCTACAC  
 AGCACATAATTCAATGCGACTTCCCTATCGCACATCTTAGACCTTATTCTCCCTCCAGC  
 ACACATCGAAGCTGCCAGCAAGCCGTTCTCACAGTCCAAGACACTGGCATGAGCGGATACA  
 TATTGAATGTATTAGAAAAATAACAAATAGGGGTTCCCGCAGCATTCCCCGAAAAGTG  
 CCACCTGAAATTGTAACGTTAATATTGTTAAAATTGCGTTAAATTGTTAAATCAG  
 CTCATTGTTAACCAATAGGCCAAATCGGAAAATCCCTATAAAATCAAAGAACATAGACCG  
 AGATAGGGTTGAGTGTGTCAGTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCC  
 AACGTCAAAGGGCGAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACCCTA  
 ATCAAGTTTTGGGTGAGGTGCCGAAAGCAGTGGCGAGAAAGGAAGGGAAAGCGAAA  
 GATTAGAGCTTGACGGGAAAGCCGGCAACGTGGCGAGAAAGGAAGGGAAAGCGAAA  
 GGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCTAACACCACACCGC  
 CGCGCTTAATGCGCCGCTACAGGGCGCGTC

**Fig. 9**

**pCMV-luc-attB (SEQ ID NO: 2)**

CTCTATCGATAGGTACCGAGCTCTACCGTGCTAGCCCTCGAGCAGGATCTACATTGAA  
TCAATATTGGCAATTAGCCATATTAGTCATTGGTTATAGCATAAATCAATATTGGCTATT  
GGCATTGCATACGTTATCTATATCATAATATGTACATTATATTGGCTATGTCCAATA  
TGACCGCCATGTTGACATTGATTATTGACTAGTTATTAGTAATCAATTACGGGTCTATT  
AGTCATAGCCCCATATGGAGTCCCGTACATAACTACGGTAAATGGCCCGCTGGCT  
GACGCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCATAGTAACGCCA  
ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCACTGGCAGT  
ACATCAAGTGTATCATATGCCAAGTCCGCCCTATTGACGTCAATGACGGTAAATGGCCCG  
CCTGGCATTATGCCAGTACATGACCTACGGACTTCCACTTGGCAGTACATCAATGGCGTGGATAGCG  
TTAGTCATCGCTATTACCATGGTGATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCG  
GTTGACTCACGGGATTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTGGCTTGGC  
ACCAAAATCAACGGACTTCCAAAATGTCGTAACAACCTCGCCCCATTGACGCAAATGGC  
GGTAGCGTGTACGGTGGAGGTCTATATAAGCAGAGCTCGTTAGTGAACCGTCAGATCGC  
CTGGAGACGCCATCCACGCTGTTGACCTCCATAGAAGACACCAGGACCGATCCAGCCTCC  
CCTCGAAGCTCGACTCTAGGGCTCGAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTAC  
TGTTGGTAAAGCCACCATGGAAGACGCCAAAAACATAAAGAAAGGCCGGCCATTCTATC  
CGCTGGAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCCTGGTT  
CCTGGAACAATTGCTTTACAGATGCACATATCGAGGTGGACATCACTACGCTGAGTACTT  
CGAAATGTCCGTTGGCAGAAGCTATGAAACGATATGGCTGAATAACAAATCACAGAA  
TCGTCGTATGCAGTGAAAACCTCTTCAATTCTTATGCCGGTGGCGCTTATTATC  
GGAGTTGCAGTTGCGCCCGGAACGACATTATAATGAACGTGAATTGCTCAACAGTATGGG  
CATTGCGAGCCTACCGTGGTGGCTTCCAAAAAGGGGTTGCAAAAAATTGAAACGTGC  
AAAAAAAGCTCCAATCATCCAAAAATTATTATCATGGATTCTAAAACGGATTACCAAGGAA  
TTTCAGTCGATGTACACGTTGTCACATCTCATCTACCTCCGGTTTAATGAATACGATT  
TGTGCCAGAGTCCTCGATAGGGACAAGACAATTGCACTGATCATGAACCTCTGGATCTA  
CTGGTCTGCCTAAAGGTGTCGCTCTGCCTCATAGAACTGCCTGCGTGAGATTCTCGCATGCC  
AGAGATCCTATTTGGCAATCAAATCATTCCGGATACTGCGATTAAAGTGTGTTCCATT  
CCATCACGGTTTGGAAATGTTACTACACTCGGATATTGATATGTGGATTGAGTCGTCT  
TAATGTATAGATTGAAGAAGAGCTGTTCTGAGGAGCCTCAGGATTACAAGATTCAAAGT  
GCGCTGCTGGTGCCAACCCTATTCTCCTCTCGCCAAAAGCACTCTGATTGACAAATACGA  
TTTATCTAATTACACGAAATTGCTCTGGTGGCGCTCCCTCTCAAGGAAGTCGGGAAAG  
CGGTTGCCAAGAGGTTCCATCTGCCAGGTATCAGGCAAGGATATGGCTCACTGAGACTACA  
TCAGCTATTCTGATTACACCGAGGGGATGATAAAACGGCGCGTGGTAAAGTTGTTCC  
ATTTTTGAAGCGAAGGTTGTGGATCTGGATACCGGGAAAACGCTGGCGTTAATCAAAGAG  
GCGAACTGTGTGAGAGGTCTATGATTATGTCCGGTTATGTAAACAAATCCGGAAGCGACC  
AACGCCTTGATTGACAAGGATGGATGGCTACATTCTGGAGACATAGCTTACTGGACGAAGA  
CGAACACTCTTCATCGTTGACCGCCTGAAGTCTCTGATTAAAGTACAAAGGCTATCAGGTGG  
CTCCCGCTGAATTGGAATCCATCTGCTCCAACACCCCAACATCTCGACGCAGGTGTCGCA  
GGTCTTCCCGACGATGACGCCGGTGAACCTCCGCCGCGTGTGTTGGAGCACGGAAA  
GACGATGACGGAAAAAGAGATCGTGGATTACGTCGCCAGTCAGTAACAACCGCGAAAAAGT  
TGCAGGGAGGAGTGTGTTGTGGACGAAGTACCGAAAGGTCTACCGGAAAAGATGCCGTGAATTCTA  
AGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGGAAAGATGCCGTGAATTCTA  
GAGTCGGGGCGGCCGCGCTTCGAGCAGACATGATAAGATACTTGATGAGTTGGACAAA  
CCACAACATAGAATGCAGTGAAAAAAATGCTTATTGTGAAATTGTGATGCTATTGCTTTA  
TTTGTAAACCATTATAAGCTGCAATAAACAGTTAACACAACAATTGCAATTCTATGTT  
TCAGGTTCAAGGGGAGGTGTGGAGGTTTAAAGCAAGTAAAACCTCTACAAATGTGGTA  
AAATCGATAAGGATCAATTGGCTTCAGGTACCGTCGACGATGTAGGTACGGTCTCGAAC  
CGCGGTGCGGGTGGCCAGGGCGTGGCTCCCGGGCGCGTACTCCACCTCACCATC  
TGGTCCATCATGATGAACGGTCGAGGTGGCGGTAGTTGATCCC GGCGAACGCGCGGGCAC

CGGGAAAGCCCTGCCCTGAAACCGCTGGCGCGGTGGTACGGTGAGCACGGACGTGCGA  
CGCGTCGGCGGGTGCAGGATACGCGGGCAGCGTCAGCGGTTCTGACGGTCACGGCGGC  
ATGTCGACAGCGAATTGATCCGTCACCGATGCCCTGAGAGCCTCAACCCAGTCAGCTC  
CTTCGGTGGCGCGGGCATGACTATCGTCGCCGACTTATGACTGTCTTATCATGC  
AACTCGTAGGACAGGTGCCGGCAGCGCTCTCCGCTCGCTACTGACTCGCTGCGCTC  
GGCGTTCGGCTGCCGGCAGCGGTATCAGCTACTCAAAGCGGTAAACCGTTATCCACAG  
AATCAGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT  
AAAAAGGCCGCGTTGCTGGCTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAA  
TCGACGCTCAAGTCAGAGGTGGCGAAACCGACAGGACTATAAGATAACCGAGCGTTCCCC  
CTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCCTGCCGTTACCGGATACCTGTCCGCC  
TTTCTCCCTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGT  
GTAGGTCGTCGCTCCAAGCTGGCTGTGCACGAACCCCCGTTAGCCGACCGCTGCG  
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATGCCACTGGCA  
GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCCTGCTACAGAGTTCTGAA  
GTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGC  
CAGTTACCTCGGAAAAAGAGTTGGTAGCTCTGATCCGGAAACAAACCACCGCTGGTAGC  
GGTGGTTTTGTTGCAAGCAGCAGATTACCGCAGAAAAAGGATCTCAAGAAGATCC  
TTGATCTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACACGTTAACGGATTGG  
TCATGAGATTATCAAAAGGATCTCACCTAGATCCTTAAATTAAAGTAAAGTTAAA  
TCAATCTAAAGTATATGAGTAAACTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC  
ACCTATCTCAGCGATCTGTCTATTGTTCATCCATAGTTGCCTGACTCCCCGTCGTGAGA  
TAACATACGATAACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAACCGCAGACCCA  
CGCTCACCGGCTCCAGATTATCAGCAATAAACAGCCAGCCAGCGGAAGGGCGAGCGCAGAAG  
TGGTCCTGCAACTTATCCGCCTCCATCCAGTCTATTAAATTGTTGCCAGGCTACAGGATCGTGGTAG  
ATCCCCCATGTTGCAAAAAAGCGGTTAGCTCCTCGGTCCGATCGTTGTCAAGAAGTA  
AGTTGGCCGCAGTGTATTCACTCATGGTTATGGCAGCACTGCATAATTCTTACTGTCATG  
CCATCCGTAAGATGCTTTCTGTGACTGGTAGTACTCAACCAAGTCATTGAGAATAGTG  
TATGCGCGACCGAGTTGCTCTGCCGGCGTCAATACGGATAATACCGGCCACATAGCA  
GAACTTAAAAGTGCATCATTGAAAACGTTCTCGGGCGAAAACCTCTCAAGGATCTTA  
CCGCTGTTGAGATCCAGTTGATGTAACCCACTCGTCACCCACTGATCTTCAGCATCTT  
TACTTCAACAGCGTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCCGAAAAAGGGAA  
TAAGGGCAGACGGAAATGTTGAATACTCATACTCTTCAATATTGAGCATT  
TATCAGGGTTATTGTCTCATGAGCGGATAACATATTGAATGTATTAGAAAAATAACAAAT  
AGGGTTCCGCGCACATTCCCCGAAAAGTGCCACCTGACCGCCCTGTAGCGCGCATTAA  
GCGCGCGGGTGTGGTGGTACCGCGAGCGTACACTTGCCAGCGCCCTAGCGCC  
GCTCCTTCGCTTCTCCCTTCTCGCCACGTTGCCGGCTTCCCCGTCAAGCTCT  
AAATCGGGGGCTCCCTTAGGTTCCGATTTAGTGCTTACGGCACCTCGACCCAAAAAC  
TTGATTAGGGTGTGGTACGTAGTGGCCATGCCCTGATAGACGGTTTTCGCCCTTG  
ACGTTGGAGTCCACGTTCTTAATAGTGGACTCTGTTCAAACACTGGAACAACACTCAACCC  
TATCTCGGTCTATTCTTGATTTATAAGGGATTTGCCGATTCGGCTATTGGTTAAAAA  
ATGAGCTGATTTAACAAAATTAAACGCAATTAAACAAATATTACGTTACAATTCC  
CATTCGCCATTCAAGGCTGCCACTGTTGGGAAGGGCGATCGTGCGGGCTCTCGCTATT  
ACGCCAGCCAAGCTACCATGATAAGTAAGTAATATTAGGTACGGGAGGTACTGGAGCGG  
CCGCAATAAAATATCTTATTTCATTACATCTGTGTGGTTTTGTGAATCGATAG  
TACTAACATACGCTCTCCATCAAAACAAAACGAAACAAACAAACTAGCAAAATAGGCTGTC  
CCCAGTGCAAGTGCAGGTGCCAGAACATT

**Fig. 10**

**pCMV-luc-attP (SEQ ID NO: 3)**

CTCTATCGATAGGTACCGAGCTCTACCGTGCTAGCCCTCGAGCAGGATCTACATTGAA  
TCAATATTGGCAATTAGCCATATTAGTCATTGGTTATATAGCATAAATCAATATTGGCTATT  
GGCCATTGCATACGTTGTATCTATATCATAATATGTACATTATATTGGCTATGTCCAATA  
TGACCGCCATGTTGACATTGATTGACTAGTTATTAATAGTAATCAATTACGGGTCTT  
AGTTCATAGCCCATAATGGAGTTCCCGTACATAACTACGGTAAATGGCCCGCTGGCT  
GACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCA  
ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCACTGGCAGT  
ACATCAAGTGTATCATATGCCAAGTCCGCCCCATTGACGTCAATGACGGTAAATGGCCCG  
CCTGGCATTATGCCAGTACATGACCTTACGGACTTCCACTTGGCAGTACATCTACGTA  
TTAGTCATCGCTATTACCATGGTGTGCGGTTGGCAGTACATCAATGGCGTGGATAGCG  
GTTGACTCACGGGATTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTGGCTTGGC  
ACCAAAATCAACGGACTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACGCAAATGGC  
GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTAGTGAACCGTCAGATCGC  
CTGGAGACGCCATCCACGCTGTTTGACCTCCATAGAAGACACCAGGACCGATCCAGCCTCC  
CCTCGAAGCTCGACTCTAGGGGCTCGAGATCTGCGATCTAAGCTTGGCATTCCGGTAC  
TGTTGGTAAAGCCACCAGGAAGACGCCAAAACATAAAGAAAGGCCGCCATTCTATC  
CGCTGGAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATGCCCTGGTT  
CCTGGAACAATTGCTTTACAGATGCACATATCGAGGTGGACATCACTTACGCTGAGTACTT  
CGAAATGTCCGTTGGCAGAACGCTATGAAACGATATGGCTGAATAACAAATCACAGAA  
TCGTCGTATGCAGTGAAAACCTCTTCAATTCTTATGCCGGTGTGGCGCTTATTATC  
GGAGTTGCAGTTGCCCGCGAACGACATTATAATGAACGTGAATTGCTCAACAGTATGGG  
CATTCGCAGCCTACCGTGGTGTGTTCCAAAAAGGGTTGCAAAAATTGAAACGTGC  
AAAAAAAGCTCCAATCATCCAAAAATTATTATCATGGATTCTAAAACGGATTACCAGGAA  
TTTCAGTCGATGTACACGTTGTCACATCTCATCTACCTCCGGTTTAATGAATAACGATT  
TGTGCCAGAGTCCTCGATAGGGACAAGACAATTGCACTGATCATGAACCTCTGGATCTA  
CTGGTCTGCCTAAAGGTGTCGCTGCCTCATAGAACTGCCTGCGTGAGATTCTCGATGCC  
AGAGATCCTATTTGGCAATCAAATCATTCCGGATACTGCGATTAAAGTGTGTTCCATT  
CCATCACGGTTTGGAATGTTACTACACTCGGATATTGATATGTGGATTTCGAGTCGTCT  
TAATGTATAGATTGAAGAAGAGCTGTTCTGAGGAGCCTCAGGATTACAAGATTCAAAGT  
GCGCTGCTGGTGCCAACCCTATTCTCCTCTCGCCAAAAGCACTCTGATTGACAAATACGA  
TTTATCTAATTACACGAAATTGCTCTGGTGGCGCTCCCTCTCTAAGGAAGTCGGGGAAG  
CGGTTGCCAAGAGGTTCCATCTGCCAGGTATCAGGCAAGGATATGGCTCACTGAGACTACA  
TCAGCTATTCTGATTACACCGAGGGGATGATAAAACGGCGCGTGGTAAAGTTGTTCC  
ATTTTTGAAGCGAAGGTTGTGGATCTGGATAACGGAAAACGCTGGCGTTAATCAAAGAG  
GCGAACTGTGTGAGAGGTCCTATGATTATGTCCGGTTATGTAACAAATCCGGAAGCGACC  
AACGCCTTGATTGACAAGGATGGATGGCTACATTCTGGAGACATAGCTACTGGACGAAGA  
CGAACACTCTTCATCGTTGACCGCCTGAAGTCTCTGATTAAGTACAAAGGCTATCAGGTGG  
CTCCCGCTGAATTGGAATCCATCTGCTCCAACACCCCAACATCTCGACGCAGGTGTCGCA  
GGTCTCCCGACGATGACGCCGGTGAACCTCCCGCCGGTTGTTGGAGCACGGAAA  
GACGATGACGGAAAAAGAGATCGTGGATTACGTCGCCAGTCAGTAACAACCGCGAAAAAGT  
TGCAGGGAGGAGTTGTGTTGTGGACGAAGTACCGAAAGGTCTTACCGAAAACCTCGACGCA  
AGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGAAAGATGCCGTGTAATTCTA  
GAGTCGGGGCGGCCGGCGCTCGAGCAGACATGATAAGATACATTGATGAGTTGGACAAA  
CCACAACTAGAATGCAGTGAACAAATGCTTATTGTGAAATTGATGCTATTGCTTTA  
TTTGTAAACCATTATAAGCTGCAATAACAAAGTTAACAAACAATTGCATTCTATTATGTT  
TCAGGTTCAGGGGGAGGTGTGGAGGTTTAAAGCAAGTAAAACCTCTACAAATGTGGTA  
AAATCGATAAGGATCAATTGGCTTCGACTAGTACTGACGGACACACCGAAGGCCCGGGC  
AACCTCAGCGGATGCCCGGGCTTCACGTTTCCAGGTCAAGAAGCGGTTTGGAGTA  
GTGCCCAACTGGGTAACCTTGAGTTCTCAGTTGGGGCGTAGGGTCGCCGACATGAC  
ACAAGGGTTGTGACCGGGGTGGACACGTACGCCGGTGTACGACCGTCAGTCGCGCGAGC

GCGACTAGTACAAGCCGAATTGATCCGTCACCGATGCCCTGAGAGCCTCAACCCAGTCA  
GCTCCTTCCGGTGGCGCGGGGATGACTATCGTCGCCGACTTATGACTGTCTTCTTATC  
ATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTCCGCTCCTCGCTCACTGACTCGCTGC  
GCTCGGTGTTGGCTGCCGAGCGGTATCAGCTCACTCAAAGGCGGTAAATACGGTTATCC  
ACAGAATCAGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA  
CCGTAAAAAGGCCGTTGCTGGCTTTCCATAGGCTCCGCCCTGACGAGCATCACA  
AAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATAACCAGGCTT  
CCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTC  
CGCCTTCTCCCTCGGAAGCGTGGCGTTCTCAATGCTCACGCTGTAGGTATCTCAGTT  
CGGTGTAGGTCGTCGCTCCAAGCTGGCTGTGCACGAACCCCCGTTAGCCGACCGC  
TGCCTTATCCGTAACATCGTCTTGAGTCCAACCCGTAAGACACGACTATGCCACT  
GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTACAGAGTTCT  
TGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCCTCTGCTG  
AAGCCAGTTACCTCGGAAAAAGAGTTGGTAGCTCTGATCCGGAAACAAACCACCGCTGG  
TAGCGGTGGTTTTGTTGCAAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAG  
ATCCTTGATCTTCTACGGGTCTGACGCTCAGTGGAACGAAAACACGTTAAGGGATT  
TTGGTCATGAGATTATCAAAAGGATCTCACCTAGATCCTTAAATTAAAAATGAAGTT  
TAAATCAATCTAAAGTATATGAGTAAACTGGTCTGACAGTTACCAATGCTTAATCAGTG  
AGGCACCTATCTCAGCGATCTGCTATTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTG  
TAGATAACTACGATAACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAACCGCGAGA  
CCACGCTACCAGCTCCAGATTATCAGCAATAAACAGCCAGCCGGAGGGCCGAGCGCA  
GAAGTGGCCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGGCCGGAGCTAGA  
GTAAGTAGTTGCCAGTTAATAGTTGCGCAACGTTGCTGCCATTGCTACAGGCATCGTGGT  
GTCACGCTCGTCTGGTATGGCTTCATTGCTCCAGTCCGGTCCCAACGATCAAGGCGAGTTA  
CATGATCCCCATGTTGCAAAAAGCGGTAGCTCCCTCGATCGTTGCTCAGA  
AGTAAGTTGCCCGCAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGT  
CATGCCATCCGTAAGATGCTTCTGTGACTGGTAGTACTCAACCAAGTCATTCTGAGAAT  
AGTGTATGCCCGACCGAGTTGCTCTGCCGGCGTCAATACGGATAATACCGGCCACAT  
AGCAGAACTTAAAGTGTCACTCATGGAAACGTTCTCGGGGAAACACTCTCAAGGAT  
CTTACCGCTGTTGAGATCCAGTCGATGTAACCCACTCGTGCACCCAACTGATCTCAGCAT  
CTTTACTTCACCAGCGTTCTGGTGAGCAAAACAGGAAGGCAAAAGGCCAAAAAG  
GGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTCCTTTCAATATTATTGAAG  
CATTATCAGGGTATTGTCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAAAC  
AAATAGGGGTTCCCGCACATTCCCCAAAAGTGCCACCTGACGCGCCCTGTAGCGCGCA  
TTAAGCGCGGGTGTGGTTACGCGCAGCGTACCGCTACACTGCCAGCGCCCTAGC  
GCCGCTCTTCCCTTCTCCCTTCTCGCCACGTTGCCGGCTTCCCGTCAAG  
CTCTAAATCGGGGCTCCCTTAGGGTTCCGATTAGTGTCTTACGGCACCTCGACCCAAA  
AAACTGATTAGGGTAGGTTCACGTAGTGGGCCATGCCCTGATAGACGGTTTTCGCC  
TTGACGTTGGAGTCCACGTTCTTAATAGTGGACTCTGTTCCAAACTGGAACAAACACTCA  
ACCCTATCTCGGTCTATTCTTGATTATAAGGGATTGCGATTTCGGCTATTGGTTA  
AAAAATGAGCTGATTAAACAAAAATTAAACGCGAATTAAACAAATATTAAACGTTACAAT  
TTCCCATTGCCATTAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCCGGCTCTCGC  
TATTACGCCAGCCAAAGCTACCATGATAAGTAAGTAATATTAAAGGTACGGGAGGTACTTGG  
GCCGGCGCAATAAAATCTTATTTCATTACATCTGTGTGGTTTGTGAATCG  
ATAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAACAAACTAGCAAAATAGGC  
TGTCCCCAGTGCAAGTGCAGGTGCCAGAACATT

**Fig. 11**

**pCMV-pur-attB (SEQ ID NO: 4)**

CTAGAGTCGGGGCGGCCGCGCTCGAGCAGACATGATAAGATACTTGTGAGTTGGAC  
AAACCACAACATAGAATGCAGTAAAAAAATGCTTATTGTGAAATTGTGATGCTATTGCT  
TTATTGTAAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAATTGCATTCTTAT  
GTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTAAAGCAAGTAAACCTCTACAAATGTG  
GTAAAATCGATAAGGATCAATTGGCTTCAGGTACCGTCGACGATGTAGGTACGGTCTCGA  
AGCCCGGGTGCAGGGTGCAGGGCGTGCCTGGCTCCCCGGCGGTACTCCACCTCACCC  
ATCTGGTCCATCATGATGAACGGGTGAGGTGGCGGTAGTTGATCCCAGGGAACGGCG  
CACCGGGAAAGCCCTCGCCCTCGAAACCGCTGGCGCGGTGGTCACGGTGAGCACGGACGTG  
CGACGGCGTCGGCGGGTGCAGGATACGCGGGCAGCGTCAGCGGTCTCGACGGTCACGGCG  
GGCATGTCGACAGCCGAATTGATCCGTGACCGATGCCCTTGAGAGCCTCAACCCAGTCAG  
CTCCTTCCGGTGGCGCGGGCATGACTATCGTCGCCGACTTATGACTGTCTTCTTATCA  
TGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTCCGCTCTCGCTACTGACTCGCTGCG  
CTCGGTCGTTGGCTGCAGCGGTATCAGCTCACTCAAAGGCGTAATACGGTTATCCA  
CAGAATCAGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC  
CGTAAAAAGGCCGCGTTGCTGGCGTTCCATAGGCTCCGCCCCCTGACGAGCATCACAA  
AAATCGACGCTCAAGTCAGAGGTGGCGAAACCGACAGGACTATAAAGATACCAGGCGTTTC  
CCCCTGGAAGCTCCCTCGCGCTCTCCTGTTCCGACCCCTGCCGTTACCGGATACCTGTCC  
GCCTTCTCCCTCGGAAGCGTGGCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTC  
GGTAGGTGCTCGCTCCAAGCTGGCTGTGACGAACCCCCCGTTAGCCGACCGCT  
GCGCCTTATCCGGTAACACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATGCCACTG  
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT  
GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGA  
AGCCAGTTACCTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGT  
AGCGGTGGTTTTGCAAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGA  
TCCTTGATCTTCTACGGGCTTGACGCTCAGTGGAACGAAAACACGTTAACGGATT  
TGGTCATGAGATTATCAAAAAGGATCTCACCTAGATCCTTAAATTAAAATGAAGTTTT  
AAATCAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA  
GGCACCTATCTCAGCGATCTGTCTATTGTTCATCCAGTTGCTGACTCCCCGTCGTGT  
AGATAACTACGATAACGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAACCGCAGAC  
CCACGCTACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCCGGAAAGGCCAGCGCAG  
AAGTGGCCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGTTGCCGGAAAGCTAGAG  
TAAGTAGTTGCCAGTTAATAGTTGCCAACGTTGCTTGCCATTGCTACAGGCATCGTGGTG  
TCACGCTCGTCTGGTATGGCTTCAATTGCTCAGCTCCGGTCCACGATCAAGGCAGTTAC  
ATGATCCCCATGTTGCAAAAAAGCGGTTAGCTCCTCGGTCCGATCGTTGTCAGAA  
GTAAGTTGCCAGTGTATCACTCATGGTTATGGCAGCAGTCATAATTCTCTACTGTC  
ATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATA  
GTGTATGCGGCCGACCGAGTTGCTCTTGGCGTCAATACGGATAATACCGGCCACATA  
GCAGAACTTAAAGTGTCTCATCATTGGAAAACGTTCTCGGGCGAAAACCTCAAGGATC  
TTACCGCTGTTGAGATCCAGTGTGATGTAACCAACTCGTCGACCCAACTGATCTCAGCATC  
TTTACTTCAACCAGCGTTCTGGGTGAGCAAAACAGGAAGGAAAATGCCGAAAAAGG  
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCAATATTGAAGC  
ATTATCAGGGTTATTGTCATGAGGGATAACATATTGAATGTATTAGAAAAATAACA  
AATAGGGGTCCGCGCACATTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGCGCAT  
TAAGCGGGCGGGTGTGGTGGTTACGCGCAGCGTACACCTGCCAGCGCCCTAGCG  
CCCGCTCCTTCGCTTCTCCCTTCTCGCCACGTTGCCGGCTTCCCGTCAAGC  
TCTAAATCGGGGCTCCCTTACGGTTGCTTACGTTAGTGCTTACGGCACCTCGACCCAAA  
AACTTGATTAGGGTGATGGTTACGTAGTGGCCATGCCCTGATAGACGGTTTCCG  
TTGACGTTGGAGTCCACGTTCTTAATAGTGGACTCTTGTGTCACGAAACACTCAA  
CCCTATCTCGGTCTATTCTTGTATTATAAGGGATTGCGATTGCGCTATTGGTTAA  
AAAATGAGCTGATTAAACAAAAATTAAACGCGAATTAAACAAAATATTAAACGTTACAATT

TCCCATTGCCATTAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGCCCTTCGCT  
ATTACGCCAGCCAAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTTGGAG  
CGGCCGCAATAAAATATCTTATTTCATTACATCTGTGTGTTTTGTGTGAATCGA  
TAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCT  
GTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTATCGATAGGTACCGAGCTCTACGC  
GTGCTAGCCCTCGAGCAGGATCTACATTGAATCAATATTGGCAATTAGCCATATTAGTCA  
TTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCAT  
AATATGTACATTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGAC  
TAGTTATTAATAGTAATCAATTACGGGTCAATTAGTCATAGCCATATATGGAGTTCCGCG  
TTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCAACGACCCCCGCCATTGACG  
TCAATAATGACGTATGTTCCATAGTAACGCCAATAGGGACTTCCATTGACGTCAATGGGT  
GGAGTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCG  
CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTA  
CGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCG  
GTTTGGCAGTACATCAATGGCGTGATAGCGGTTGACTCACGGGGATTCCAAGTCTCC  
ACCCCATTGACGTCAATGGAGTTGGCACCAAAATCAACGGACTTCCAAAATGT  
CGTAACAACCTCCGCCATTGACGCAAATGGCGGTAGGCGTGTACGGTGGGAGGTCTATAT  
AAGCAGAGCTCGTTAGTGAACCCTCAGATCGCCTGGAGACGCCATCCACGCTGTTGACC  
TCCATAGAAGACACCGGACCGATCCAGCCTCCCTCGAACGCTCGACTCTAGGGCTCGAGA  
TCTGCGATCTAAGTAAGCTTGCATGCCTGCAGGTGGCCACGACCGGTGCCGCCACCAT  
CCCCTGACCCACGCCCTGACCCCTCACAAGGAGACGCCATTCCATGACCGAGTACAAGCCC  
ACGGTGCCTCGCCACCCCGGACGACGTCCCCCGGGCGTACGCACCCCTCGCCGCCGTT  
CGCCGACTACCCGCCACGCCACACCGTCGACCCGGACCGCCACATCGAGCGGGTCACCG  
AGCTGCAAGAACTCTTCCTCACCGCGTCGGCTCGACATCGGCAAGGTGTGGGTCGCGGAC  
GACGGCGCCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCAAGCGGGGGCGGTGTCG  
CGAGATCGGCCCGCGCATGGCGAGTTGAGCGGTTCCCGCTGGCCGCGCAGCAACAGATGG  
AAGGCCTCTGGCGCCGACCGGCCAAGGAGCCCGTGGCTCTGGCCACCGTCGGCGTC  
TCGCCGACCACCAAGGGCAAGGGTCTGGGAGCGCCGTGCTCCCCGGAGTGGAGGCGGC  
CGAGCGCGCCGGGTGCCCTCCTGGAGACCTCCGCCCCGCAACCTCCCTCTACG  
AGCGGCTCGGCTTCACCGTCACCGCCACGTCGAGGTGCCGAAGGACCGCGCACCTGGTGC  
ATGACCCGCAAGCCCAGTGCCTGACGCCGCCCCACGACCCGCAGCGCCGACCGAAAGGAG  
CGCACGACCCATGGCTCCGACCGAAGCCGACCCGGGCGGCCGACCCGCACCCGCC  
CCCGAGGCCACCGACT

**Fig. 12**

**pCMV-pur-attP (SEQ ID NO: 5)**

CTAGAGTCGGGGCGGCCGGCGCTTCGAGCAGACATGATAAGATAACATTGATGAGTTGGAC  
AAACCACAACATAGAATGCAGTGAACAAATGCTTATTGAAATTGATGCTATTGCT  
TTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAATTGCATTCAATTAT  
GTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTAAAGCAAGTAAAACCTCTACAAATGTG  
GTAAAATCGATAAGGATCAATTGGCTTCGACTAGTACTGACGGACACACCGAAGCCCCGGC  
GGCAACCCTCAGCGGATGCCCGGGGCTTCACGTTCCCAGGTAGAACAGCGGTTTCGGGA  
GTAGTGCCCCAAGCTGGGTAACCTTGAGTTCTCTCAGTTGGGGCGTAGGGTCGCCGACAT  
GACACAAGGGGTTGTGACCGGGTGGACACGTACGCGGGTGTACGACCGTCAGTCGCGCG  
AGCGCGACTAGTACAAGCGAATTGATCCGTCGACCGATGCCCTGAGAGCCTCAACCCAG  
TCAGCTCCTCCGGTGGCGCGGGCATGACTATCGTCGCCGACTTATGACTGTCTTCTT  
ATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTTCCGCTCGCTCACTGACTCGC  
TGCCTCGGTCGTTGGCTGCGAGCGGTATCAGCTCACTCAAAGGCGTAATACGTTA  
TCCACAGAATCAGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG  
GAACCGTAAAAGGCCGCGTTGGCTGGCTTTCCATAGGCTCCGCCCTGACGAGCATT  
ACAAAAATCGACGCTCAAGTCAGAGGTGGCAAACCCGACAGGACTATAAACGATACCGCG  
TTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCCTGCCGTTACCGGATACCT  
GTCCGCCTTCTCCCTTGGAAAGCGTGGCGCTTCTCAATGCTCACGCTGTAGGTATCTCA  
GTTGGTGTAGGTCGTTGGCTCCAAGCTGGCTGTGCACGAACCCCCGTTAGCCCGAC  
CGCTGCGCCTTATCCGTAACATCGTCTTGAGTCCAACCCGTAAGACACGACTTATGCC  
ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT  
TCTTGAAAGTGGTGGCCTAACTACGGCTACACTAGAACAGTATTGGTATCTGCGCTCTG  
CTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTGATCCGGCAAACAAACCACCGC  
TGGTAGCGGTGGTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG  
AAGATCCTTGATCTTCTACGGGTCTGACGCTCAGTGGAACGAAAACACGTTAACGG  
ATTTGGTCATGAGATTATCAAAAGGATCTCACCTAGATCCTTAAATTAAAAATGAAG  
TTTAAATCAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCA  
GTGAGGCACCTATCTCAGCGATCTGTCTATTGTTCATCCATAGTTGCCTGACTCCCCGTC  
GTGTAGATAACTACGATAACGGGAGGGCTTACCATCTGGCCCGAGTGCTGCAATGATAACCGCG  
AGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACAGCCAGCCGGAAAGGGCCGAGC  
GCAGAAAGTGGCCTGCAACTTATCCGCCTCCATCCAGTCTATTAAATTGTTGCCGGAAAGCT  
AGAGTAAGTAGTTGCCAGTTAATAGTTGCGCAACGTTGCTGCTACAGGCATCGT  
GGTGTACGCTCGTGGTATGGCTTCATTCTAGCTCCGGTCCACGATCAAGGCGAG  
TTACATGATCCCCATGTTGCAAAAAAGCGGTTAGCTCCTCGGTCTCGATCGTTGTC  
AGAAGTAAGTTGGCGCAGTGTATCACTCATGGTTATGGCAGCAGTCATAATTCTTAC  
TGTCTGCCATCCGTAAGATGCTTTCTGACTGGTGAGTACTCAACCAAGTCATTGAG  
AATAGTGTATGCGCGACCGAGTTGCTCTGCCCGGTCAATAACGGGATAATACCGCGCCA  
CATAGCAGAACCTTAAAGTGTCTCATATTGGAAAAACGTTCTCGGGCGAAAACCTCTCAAG  
GATCTTACCGCTGGTGGAGATCCAGTTGATGTAACCCACTCGTGCACCCAACTGATCTTCA  
CATCTTTACTTCAACCAGCGTTCTGGGTGAGCAAAAACAGGAAGGAAAATGCCGAAAA  
AAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTCCTTTCAATATTATTG  
AAGCATTATCAGGGTTATTGTCATGAGCGGATAACATATTGAATGTATTAGAAAAATA  
AACAAATAGGGTTCCGCGCACATTCCCCAAAAGTGCCACCTGACGCGCCCTGAGCGG  
GCATTAAGCGCGGGGGTGGTGGTACGCGCAGCGTACCGCTACACTGCCAGCGCCCT  
AGCGCCCGCTCTTCTGCTTCTCCCTTCTCGCCACGTTGCCGGCTTCCCCGTC  
AAGCTCTAAATCGGGGGCTCCCTTGTAGGGTCCGATTAGTGTCTTACGGCACCTCGACCC  
AAAAAAACTTGATTAGGGTGTGGTACGTTACGTAGTGGGCCATGCCCTGATAGACGGTTTC  
CCCTTGACGTTGGAGTCCACGTTCTTAATAGTGGACTCTGTTCCAAACTGGAACACAC  
TCAACCCTATCTCGGTCTATTCTTTGATTATAAGGGATTGCGCATTGCCCTATTGG  
TTAAAAAAATGAGCTGATTAAACAAAAATTAAACGCGAATTAAACAAAATATTACGTTAC  
AATTCCCATTGCCATTAGGCTGCGCACTGTTGGGAAGGGCGATCGGTGCGGGCTCTT

CGCTATTACGCCAGCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTT  
GGAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTGTGTGAA  
TCGATAGTACTAACATACGCTCTCCATCAAAACAAAAGAAACAAAACAAACTAGCAAAATA  
GGCTGTCCCCAGTGCAAGTGCAAGGTGCCAGAACATTCTATCGATAGGTACCGAGCTCTT  
ACGCGTGCTAGCCCTCGAGCAGGATCTACATTGAATCAATATTGGCAATTAGCCATATTA  
GTCATTGGTTATATAGCATAAAATCAATATTGGCTATTGGCATTGCATACTGTTATCTATA  
TCATAATATGTACATTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTAT  
TGACTAGTTATTAATAGTAATCAATTACGGGTCTAGTTCATAGCCCATAATGGAGTTC  
CGCGTTACATAACTACGGTAAATGGCCCGCTGGCTGACCGCCAACGACCCCCGCCATT  
GACGTCAATAATGACGTATGTTCCATAGTAACGCCAATAGGGACTTCATTGACGTCAAT  
GGGTGGAGTATTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGT  
CCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAGTACATGAC  
CTTACGGGACTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGA  
TGCGGTTTGGCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGGATTCCAAGT  
CTCCACCCCATTGACGTCAATGGGAGTTGGCACC AAAATCAACGGGACTTCACAA  
ATGTCGTAACAACTCCGCCATTGACGCAAATGGCGGTAGGCGTGTACGGTGGGAGGTCT  
ATATAAGCAGAGCTCGTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTT  
GACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCCTCGAAGCTGACTCTAGGGCTC  
GAGATCTCGATCTAAGTAAGCTTGCATGCCTGCAGGTGCCACGACCGGTGCC  
CCATCCCTGACCCACGCCCTGACCCCTCACAGGAGACCTCCATGACCGAGTACAA  
GCCCACGGTGCCTGCCACCGCGACGACGTCCCCGGCGTACGCACCCCTGCC  
CGTTGCCGACTACCCGCCACCGGCCACACCGTCGACCCGGACCGCCACATCGAGCGGGTC  
ACCGAGCTGCAAGAACTCTCCTCACCGCGTCGGGCTCGACATCGCAAGGTGTGGGTC  
GGACGACGGCGCCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCAAGCGGGGGCGGTGT  
TCGCCGAGATCGGCCCGCGCATGGCGAGTTGAGCGGTTCCCGCTGGCCCGCAGCAACAG  
ATGGAAGGCCTCTGGCGCCGACCGGCCAACGGAGCCCGTGGTCTGGCCACCGTC  
CGTCTGCCGACCAACCAGGGCAAGGGCTGGCAGCGCCGTGCTCCCCGGAGTGGAGG  
CGGCCGAGCGCGCCGGGTGCCCTCCTGGAGACCTCCCGCCCCGCAACCTCCCCTC  
TACGAGCGGCTCGGCTTCACCGTCACCGCCACGTCGAGGTGCCGAAGGACCGCG  
GTGCATGACCCGCAAGCCGGTGCCTGACGCCGCCCCACGACCCGAGCGCCGACCGAAA  
GGAGCGCACGACCCATGGCTCCGACCGAAGCCGACCCGGGGCCCCGCCACCCGCACC  
CGCCCCCGAGGCCACCGACT

*Fig. 13*

**pCMV-EGFP-attB (SEQ ID NO: 6)**

CTAGAGTCGGGGCGGCCGCGCTCGAGCAGACATGATAAGATAACATTGATGAGTTGGAC  
AAACCACAACATAGAATGCAGTAAAAAAATGCTTATTTGTGAAATTGTGATGCTATTGCT  
TTATTTGTAACCATTATAAGCTGCAATAAACAGTTAACACAACAATTGCATTCAATTAT  
GTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTAAAGCAAGTAAAACCTCTACAAATGTG  
GTAAAATCGATAAGGATCAATTGGCTTCAGGTACCGTCGACGATGTAGGTACCGTCTCGA  
AGCCGCGGTGCGGGTGCAGGGCGTGCCTGGGCTCCCCGGCGCGTACTCCACCTCACCC  
ATCTGGTCATCATGATGAACGGTCGAGGTGGCGGTAGTTGATCCGGCGAACGGCGGGCG  
CACCGGGAAGCCCTGCCCTCGAAACCGCTGGCGCGTGGTACGGTGAGCACGGACGTG  
CGACGGCGTCGGCGGGTGCAGGATACGGGGCAGCGTCAGGGTTCTGACGGTCACGGCG  
GGCATGTGACAGCGAATTGATCCGTCGACCGATGCCCTGAGAGCCTCAACCCAGTCAG  
CTCCTTCCGGTGGCGCGGGGATGACTATCGTCGCCGACTTATGACTGTCTCTTATCA  
TGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTCCGCTCTCGCTACTGACTCGCTGCG  
CTCGGTGTTGGCTGCCGAGCGGTATCAGCTCACTCAAAGGCGTAATACGGTTATCCA  
CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC  
CGTAAAAAGGCCGCGTTGCTGGCTTTCCATAGGCTCCGCCCCCTGACGAGCATTACAA  
AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCAGGCGTTTC  
CCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCC  
GCCTTCTCCCTCGGGAAGCGTGGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTC  
GGTAGGTGCTCCAGCTGGCTGTGCAACGACCCCCCGTTAGCCGACCGCT  
GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATGCCACTG  
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTT  
GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGA  
AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTTGTACCGGAAACAAACCACCGCTGGT  
AGCGGTGGTTTTGCAAGCAGATTACCGCAGAAAAAGGATCTCAAGAAGA  
TCCTTGATCTTCTACGGGCTGACGCTCAGTGGAACGAAAACACGTTAACGTTAACGTTAACGGATT  
TGGTCATGAGATTATCAAAAGGATCTCACCTAGATCCTTAAATTAAAAATGAAGTTT  
AAATCAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA  
GGCACCTATCTCAGCGATCTGTCTATTGTTCATCCATAGTTGCCGACTCCCCGCGTGT  
AGATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAACCGCAGAC  
CCACGCTACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCCGGAAAGGGCCAGCGCAG  
AAGTGGTCTGCAACTTATCCGCCTCCATCCAGTCTATTAAATTGTTGCCGGAAAGCTAGAG  
TAAGTAGTTGCCAGTTAATAGTTGCGCAACGTTGTCATTGCTACAGGCATCGTGGTG  
TCACGCTCGTGTGGTATGGCTTCATTCAAGCTCCGGTCCAAACGATCAAGGCAGTTAC  
ATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTCGGTCCGATCGTGTCAAGAA  
GTAAGTTGCCAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTTACTGTC  
ATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATA  
GTGTATGCCGACCGAGTTGCTCTTGGCTCAATACGGATAATACCGGCCACATA  
GCAGAACTTAAAGTGCCTCATCATTGGAAAACGTTCTCGGGGGAAAACCTCAAGGATC  
TTACCGCTGAGATCCAGTTGATGTAACCCACTCGTGACCCAACTGATCTCAGCATT  
TTTACTTCAACCAGCGTTCTGGTGAGCAAAACAGGAAGGCAAAATGCCAAAAAGG  
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCAATATTATTGAAGC  
ATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAAAC  
AATAGGGTTCCGCGCACATTCCCCGAAAAGTGCCACCTGACGCCCTGTAGGGCGCAT  
TAAGCGCGGGGTGTTACGCGCAGCGTACCGCTACACTGCCAGCGCCCTAGCG  
CCCGCTCCTTCGCTTCTCCCTTCTCGCCACGTTGCCGGCTTCCCGTCAAGC  
TCTAAATCGGGGCTCCCTTACGGTTCCGATTAGTGCTTACGGCACCTCGACCCAAAA  
AACTGATTAGGGTGATGGTTCACGTAGTGGCCATGCCCTGATAGACGGTTTCGCCC  
TTGACGTTGGAGTCCACGTTAATAGTGGACTCTGTTCAAACGGAAACAACACTCAA  
CCCTATCTCGGTCTATTCTTTGATTATAAGGGATTGCGATTCGCGCTATTGGTTAA  
AAAATGAGCTGATTAACAAAAATTAAACGCGAATTAAACAAATATTACGTTACAATT

TCCCATTGCCATTAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGGGGCCTTCGCT  
ATTACGCCAGCCAAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTGGAG  
CGGCCGCAATAAAATATCTTATTTCATTACATCTGTGTGGTTTTGTGTGAATCGA  
TAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCT  
GTCCCCAGTGCAAGTGCAGGTGCCAGAACATTCTATCGATAGGTACCGAGCTTACGC  
GTGCTAGCCCTCGAGCAGGATCTACATTGAATCAATATTGGCAATTAGCCATATTAGTCA  
TTGGTTATATAGCATAAATCAATATTGGCTATTGCCATTGCATACGTTGTATCTATATCAT  
AATATGTACATTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGAC  
TAGTTATTAATAGTAATCAATTACGGGTCAATTAGTCATAGCCCATAATGGAGTTCCGCG  
TTACATAACTACGGTAAATGGCCCGCCTGGCTGACGCCAACGACCCCCGCCATTGACG  
TCAATAATGACGTATGTTCCCAGTAGTAACGCCAATAGGGACTTCCATTGACGTCAATGGGT  
GGAGTATTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC  
CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTA  
CGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG  
GTTTGGCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGGATTCCAAGTCTCC  
ACCCCATTGACGTCAATGGAGTTGGCACCAAAATCAACGGACTTCCAAAATGT  
CGTAACAACCTCCGGCCATTGACGCAAATGGCGGTAGGCCTGTACGGTGGAGGTCTATAT  
AAGCAGAGCTCGTTAGTGAACCGTCAGATGCCCTGGAGACGCCATCCACGCTGTTGACC  
TCCATAGAAGACACCGGGACCGATCCAGCCTCCCTGAAGCTCGACTCTAGGGCTCGAGA  
TCCCCGGGTACCGGTGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGTGGTGC  
CCATCCTGGTCGAGCTGGACGGCGACGTAAACGCCACAAGTTCAGCGTCCGGCGAGGGC  
GAGGGCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATCTGCACCAACGGCAAGCTGCC  
CGTGCCCTGGCCCACCCCTCGTGAACCGACCTACGGCGTGCAGTGCCTCAGCCGCTACC  
CCGACCATGAAAGCAGCACGACTTCTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAG  
CGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCCGCCGAGGTGAAGTTCGAGGG  
CGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCC  
TGGGGCACAAGCTGGAGTACAACACAGCCACAACGTCTATATCATGGCCGACAAGCAG  
AAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT  
CGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCCGACAACC  
ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC  
CTGCTGGAGTTCGTGACCGCCGCCGGATCACTCTGGCATGGACGAGCTGTACAAGTAAAG  
CGGCCGCTCGAGCATGCAT

*Fig. 14*

**p-12.0-lys-LSP1FNMM-CMV-pur-attB (SEQ ID NO: 7)**

GGGCTGCAGGAATTGCGATTGCCGCCTCTTGATATTCACTCTGTGATTTCATCTCTTCT  
TGCCGATGAAAGGATATAACAGTCTGTATAACAGTCTGTGAGGAAATACTGGTATTC  
TGATCAGTGTGTTATAAGTAATGTTGAATATTGGATAAGGCTGTGTCCTTGTCTGGG  
AGACAAAAGCCCACAGCAGGTGGTGGGGTGGCAGCTCAGTGACAGGAGAGGTTTT  
TTGCCTGTTTTTTTTTTTTTTAAGTAAGGTGTTCTTTCTTAGTAAATT  
CTACTGGACTGTATGTTGACAGGTGAGAACATTCTCAAAAGAACCTTGAA  
CTGTACAGCCCTTTCATTCCCTTTGCTTCTGTGCCAATGCCCTGGTTCTGATT  
GCATTATGGAAAACGTTGATCGGAACGTGAGGTTTATTATAGTGTGGCTGAAAGCTG  
GATAGCTGTTGTTACACGAGATACTTATTAGTTAGGCCAGCTGATGCTTATT  
CCTTGAAGTAGTGAGCGTTCTGGTTTCCCTTGAACACTGGTGAGGCTTAGATT  
CTAATGGGATTTTACCTGATGATCTAGTTGCATACCCAAATGCTGTAAATGTT  
GTTAACATGTTGATAACTCGGATTACATGTTGATATACTGTCATCTGTGTT  
AAAATATATGGCATTAGAAATACGTAATTCTGATTTCTTTTATCTCTATGCT  
CTGTGTGTACAGGTCAAACAGACTTCACTCCTATTATTATAGAATT  
GTCGTTGGTTCTGTGTAAGGATACAGCCTAAATTCTAGAGCGATGCTCAGTAAG  
GCGGGTTGTCACATGGGTTCAAATGTAACACGGGACGTTGGCTGCTGCCCTCCGAGATC  
CAGGACACTAAACTGCTCTGCACTGAGGTATAAACGCTTCAGATCCCAGGAAAGTGCAGA  
TCCACGTGCATATTCTAAAGAAGAATGAATACCTTCTAAATATTGGCATAGGAAGCAA  
GCTGCATGGATTGTTGGACTAAATTATTGGTAACGGAGTGCATAGGTTAAACAC  
AGTTGCAGCATGCTAACGAGTCACAGCGTTATGCAGAAGTGATGCC  
GCTGTTACGGCACTGCCCTGCAGTGAGCATTGCAGATAGGGTGGGCTTGTGCGT  
TTCCACACGCTGCCACACGCCACCTCCGGAACACATCTCACCTGCTGGTACT  
ACCATCTTAGCAGTAGTGTACTATGAAACAGAGAAGTCCCTCAGTTGGATATT  
CATGGGATGCTTTCCATGTTGGCAAAGTATGATAAAGCATCTATTGTA  
TGCACCTGTTAGTCCTGAATCCTTCTATAGCACCCTATTGCAGCAGGTGAGGCT  
GTGTTGGCTGTGCTGCTCAATCTTAAAGCTCTTGGAAATA  
AAGTCTCTGAAGATAGTAAACAGTACTTACCTTGA  
TCCGGCTGCTGCTGCTCAATCTTAAAGCTCTTGGAAATA  
ACACTGACTGATTG  
AAGTCTCTGAAGATAGTAAACAGTACTTACCTTGA  
TCCGGCTGCTGCTGCTCAATCTTAAAGCTCTTGGAAATA  
ACACTGACTGATTG  
TGTAAAAGAATTCCGCCTATTCA  
TTGGAATATTCAAGTAATAGACTTGGCCTCAC  
AAAATATTGAAAGAGACATTCTGCTGATCTCA  
AATGTAAGAAAATGAGGAGTGC  
GTGCTTAAAGAATGGCTTACAAGT  
ACATTCC  
ACATAAGGCTGTATTGCTCCTGTACCA  
GTCTGGGTAAACTATTCAAGAAATGGCTTGA  
ATGCAGAGTTGC  
GATTTATATGTGTATAGTAAGCAG  
TTCTGATTCAACAAGGTAAG  
TTCATTGACTGTCTGATATCCTGCAG  
ACATCAAAACTTAACGTGAGCTCAG  
TAGTTAGATGTTACTTGA  
CCTTTAGGAGACTGTTAAGCA  
TAGTTACCTGTATTGA  
TTCATTGACTGTCTGATATCCTGCAG  
ACATCAAAACTTAACGTGAGCTCAG  
TAGTGAGAAATGC  
TTCAATTGGAAATATTGTTCTAGACTGC  
TGGCATGA  
AAGCAAGGGCACAGGTCC  
CTTGGCTAGTGTCTAAATCTG  
GGGCTCC

AGCTTTAATGCTTCAAATTGAAGGTGGCAGGCAACTGGCCACTGGTTATTAATGCATTA  
TGTCTCAGTTCGCAGCTAACCTGGCTCTCACTATTGAGCATGGACTATAGCCTGGCTTC  
AGAGGCCAGGTGAAGGTGGATGGGTGGAAGGAGTGCTGGCTGGCTGGGGACTGTG  
GGGACTCCAAGCTGAGCTGGGGTGGCAGCACAGGGAAAGTGTGGTAACTATTTTAAG  
TACTGTGTTGCAAACGTCTCATCTGCAAATACGTAGGGTGTACTCTCGAAGATTAACAGT  
GTGGGTTCAAGTAATATATGGATGAATTCACAGTGGAAGCATTCAAGGGTAGATCATCTAACG  
ACACCAGATCATCAAGCTATGATTGGAAGCGGTATCAGAAGAGCGAGGAAGGTAAGCAGTCT  
TCATATGTTTCCCTCCACGTAAAGCAGTCTGGAAAGTAGCACCCCTGAGCAGAGACAAG  
GAAATAATTCAAGGAGCATGTGCTAGGAGAACCTTCTGCTGAATTCTACTTGCAAGAGCTT  
GATGCCTGGCTCTGGTGCCTCTGCAGCACCTGCAAGGCCAGAGCCTGTGGTGAGCTGGA  
GGGAAAGATTCTGCTCAAGTCCAAGCTCAGCAGGTCAATTGTCTTGCTCTCCCCCAGCA  
CTGTGCAGCAGAGTGGAACTGATGTCAGGCCTCCTGCAACTACCTGTTGCTGCAGGCAGA  
CTGCTCTCAGAAAAAGAGAGCTAACTCTATGCCATAGTGTGAGGTAAAATGGGTTTAAAAA  
AAGAAAACACAAAGGAAAACCGGCTGCCCATGAGAAGAAAGCAGTGGTAAACATGGTAGA  
AAAGGTGCAGAAGCCCCCAGGCAGTGTGACAGGCCCTCCTGCCACCTAGAGGCCGGAAACAA  
GCTTCCCTGCCTAGGGCTCTGCCCGCAAGTGCCTGTTCTTGGTGGTTTGGCGT  
TTGGGTTTGAGATTAGACACAAGGAAAGCCTGAAAGGAGGTGTTGGCACTATTTGGTTT  
GTAAAGCCTGTACTTCAAATATATATTGTGAGGGAGTGTAGCGAATTGGCAATTAA  
TAAAGTTGCAAGAGATTGAAGGCTGAGTAGTTGAGAGGGTAACACGTTAATGAGATCTTCT  
GAAACTACTGCTTCTAAACACTGTTGAGTGGTGAGACCTGGATAGGTGAGTGCTCTTGT  
TACATGTCTGATGCACCTGCTTGCTCTTCCATCCACATCCATGCATTCCACATCCACGCA  
TTTGTCACTTATCCATATCTGTCAATCTGACATAACCTGCTCTCGTCACTGGTCAGAA  
GAAACAGATGTGATAATCCCCAGCCGCCAAGTTGAGAAGATGGCAGTTGCTCTTCCC  
TTTTCTGCTAAGTAAGGATTCTCCTGGCTTGACACCTCACGAAATAGTCTCCTG  
TTACATTCTGGCATTATTCAAATATCTTGGAGTGCCTGCTCAAGTTGTGCTTCC  
TACTCTTAGAGTGAATGCTCTAGAGTGAAAGAGAAGGAAGAGAAGATGTTGGCCGAGTTC  
TCTGATGAACACACCTCTGAATAATGCCAAAGGTGGTTCTGAGGAACGGCAG  
CGTTGCCTCTGAAAGCAAGGAGCTGCGGAGTTGCAGTTTTGCAACTGATGGTGGAA  
CTGGTGCTTAAAGCAGATTCCCTAGGTTCCCTGCTACTCTTCTTCTGGCAGTCAGTT  
TATTCTGACAGACAAACGCCACCCCCACTGCAGGCTTAGAAAGTATGTGGCTCTGCCTGG  
GTGTGTTACAGCTCTGCCCTGGTAAAGGGATTAAAACGGGCACCATTCACTCCAAACAGG  
ATCCTCATTCACTGGATCAAGCTGTAAGGAACCTGGCTCAACCTCAAACATTAAATTGGAG  
TACGAATGTAATTAAAACGCTCTCGCATTCTCAAGTCATTAGTCTGGACTCTGCAGCA  
TGTAGGTCGGCAGCTCCACTTCTCAAAGACCACTGATGGAGGAGTAGTAAAATGGAGAC  
CGATTCAGAACACCAACGGAGTGTGCGAAGAAACTGATGGAAATAATGCATGAATTGTG  
TGGTGGACATTTTAAATACATAAAACTACTTCAAATGAGGTGGAGAAGGTCACTGTT  
TATTAGCAGCCATAAAACCAAGGTGAGCGAGTACCACTTCTCTACAAGAAAAACGATTCTG  
AGCTCTGCGTAAGTATAAGTTCTCCATAGCGGCTGAAGCTCCCCCTGGCTGCCATCT  
CAGCTGGAGTGCAGTGCCTTCTCAGCAGTAAATGGACAATACTTC  
ACAAAAAATTCTTCTTCTGTCACTGGGATCCCTACTGTGCCCTGGTTTACGTTA  
CCCCCTGACTGTTCCATTCACTGGTTGGAAAGAGAAAAAGAATTGGAAATAACATGTC  
TACGTTATCACCTCCAGCATTGGTTTAATTATGTCATAACTGGCTTAGATTGG  
AAATGAGAGGGGGTTGGGTGATTACCGAGGAACAAAGGAAGGCTTATATAAAACTCAAGTCT  
TTTATTAGAGAACTGGCAAGCTGTCAAAAACAAAAAGGCCTTACCAACAAATTAAAGTGAAT  
AGCCGCTATAGCCAGCAGGGCCAGCACGAGGGATGGTGCAGTGCTGGCACTATGCCACGGCC  
TGCTTGTGACTCTGAGAGCAACTGCTTGGAAATGACAGCAGTGGCAATTCCCTTGT  
TCAGAATGCGTAGAGCGTGTGCTTGGCGACAGTTTTCTAGTTAGGCCACTTCTTTCT  
TCTCTCCTCATTCTCCTAAGCATGTCTCCATGCTGGTAATCCAGTCAAGTGAACGTTCAA  
CAATGAATCCACTGTAGGATTCTCGTGGTGAAGATCAAATCTTGTGAGGTCTATAAAAT  
ATGGAAGCTTATTATTTCGTTCTCCATATCAGTCTCTATGACAATTACACATCCAC  
CACAGCAAATTAAAGGTGAAGGAGGCTGGTGGATGAAGAGGGCTTAGCTTACGTTCT  
TCCTGCAAGGCCACAGGAAATGCTGAGAGCTGAGAATACAGCCTGGGTAAGAAGTTCA

GTCTCCTGCTGGACAGCTAACGCATCTTATAACCCCTTGAGACTCATCTTAGGACCAA  
ATAGGGTCTATCTGGGTTTGTCTGCTGTTCTCCTGGAAAGGCTATCTCACTATTCA  
CTGCTCCCACGGTACAAACCAAAGATAACAGCCTGAATTCTAGGCCACATTACATAAA  
TTTGACCTGGTACCAATATTGTTCTATATAGTTATTCCTTCCCCACTGTGTTAACCCC  
TTAAGGCATTAGAACAACTAGAACATAGAATGGTTGGATTGGAAGGGGCCTAAACATC  
ATCCATTCCAACCCCTGCCCAGGGCTGCTGCCACCCACTGGCTCAGGCTGCCAGGGCC  
CCATCCAGCCTGGCCTGAGCACCTCCAGGGATGGGGCACCCACAGCTCTGGCAGCCT  
GTGCCAACACCTCACCCTCTGGTAAAGAATTCTCTTTAACATCTAAATCTCT  
TCTCTTTAGTTAAAGCCATTCCCTTTCCCCTGCTATCTGTCCAAGAAATGTGTATT  
GGTCTCCCTCCTGCTTATAAGCAGGAAGTACTGGAAGGCTGCAGTGAGGTCTCCCCACAGCC  
TTCTCTTCTCCAGGCTGAACAAAGCCCAGCTCCTCAGCCTGTCTCGTAGGAGATCATCTT  
GTGCCCTCCTCTGGACCCATTCCAACAGTTCCACGGCTTCTTGTGGAGCCCCAGGTCTGG  
ATGCAGTACTTCAGATGGGCCTTACAAAGGCAGAGCAGATGGGACAATCGCTTACCCCTC  
CCTGCTGGCTGCCCTGTTGATGCAGCCCAGGGTACTGTTGGCCTTCAGGCTCCCAGAC  
CCCTTGCTGATTGTGTCAGCTTCCACCAAGAACCCACGCTTGGTTAACACTTC  
TGCCCTCACTCTGTAAGCTTCAAGGAGACTTCCATTCTTAGGACAGACTGTGTTACA  
CCTACCTGCCCTATTCTGCATATATACATTCAAGTTCATGTTCTGTAACAGGACAGAAT  
ATGTATTCCCTCTAACAAAAATACATGCAGAATTCTAGTGCCTACAGTAGGTTTCTG  
GCAGTATTAGCACATAGTCAATTGCTGCAAGTACCTTCCAAGCTGCCCTCCCATAAATC  
CTGTATTGGGATCAGTTACCTTGGGTAAGCTTGTATCTGCAGAGACCCCTGGGGGTT  
CTGATGTGCTTCAGCTCTGCTCTGTTCTGACTGCACCATTCTAGATCACCCAGTTGTTCC  
TGTACAACCTCCTGTCCTCCATCCTTCCCAGCTGTATCTTGACAAATACAGGCCTATT  
TTTGTGTTGCTTCAGCAGCATTAAATTCTCAGTGTCAATTGCTTCTGTTGATGCCACTG  
GAACAGGATTTCAGCAGTCTGCAAAGAACATCTAGCTGAAAACCTTCTGCCATTCAATAT  
TCTTACCAAGTTCTGTTGAGGTGAGCCATAAATTACTAGAACACTCGTCACTGACAAGT  
TTATGCATTTATTACTCTATTATGTACTTACTTGACATAACACAGACACGCACATATT  
TGCTGGGATTCCACAGTGTCTGTGTCCTCACATGGTTACTGTCATACTCCGTTAT  
AACCTTGGCAATCTGCCAGCTGCCATCACAAAGAAAAGAGATTCTTTTATTACTTCTC  
TTCAGCCAATAACAAAATGTGAGAAGCCAAACAAGAACCTGTGGGGCAGGCTGCCATCAA  
GGGAGAGACAGCTGAAGGGTTGTGAGCTCAATAGAATTAAAGAAATAAAAGCTGTGTCAG  
ACAGTTTGCCTGATTATACAGGCACGCCAACGCCAGAGAGAGGCTGTGCCAAGGCCACC  
TTGCAGTCCTGGTTGTAAGATAAGTCATAGGTAACCTTCTGGTGAATTGCGTGGAGAAT  
CATGATGGCAGTTCTGCTGTTACTATGGTAAGATGCTAAAATAGGAGACAGCAAAGTAAC  
ACTTGCTGCTGTAGGTGCTCTGCTATCCAGACAGCGATGGCACTCGCACACCAAGATGAGGG  
ATGCTCCCAGCTGACGGATGCTGGGGCAGTAACAGTGGTCCCCTGCTGCCCTGCTCATTAGC  
ATCACCTCAGCCCTCACCAAGCCATCAGAAGGATCATCCCAAGCTGAGGAAAGTTGCTCATT  
TTCTCACATCATCAAACCTTGGCCTGACTGATGCCCTCCGGATGCTTAAATGTGGTCACT  
GACATCTTATTCTATGATTCAAGTCAGAACCTCCGGATCAGGAGGGAACACATAGTG  
GGAATGTACCCTCAGCTCCAAGGCCAGATCTCCTCAATGATCATGCATGCTACTTAGGAA  
GGTGTGTGTGTAATGAGAATTGCTTTGTTCTTCTGCTGTGAGGAACATT  
TTGAATACCAGAGAAAAAGAAAAGTGTCTTCTGGCATGGGAGGAGTTGTCACACTTGCAA  
AATAAAGGATGCAGTCCAAATGTTCATATCTCAGGGCTGAGGAGGATCAGAAACTGTG  
TATACAATTCTCAGGCTCTGAAATGCAAGCTTTGAAAGCTGTTCTGGCCGAGGAGTACT  
AGTCAGAACCCCTCGGAAACAGGAACAAATGTCCTCAAGGTGCAGCAGGAGGAAACACCTTGC  
CCATCATGAAAGTGAATAACCACGCCGCTGAAGGAATCCAGCTCCTGTTGAGCAGGTGCT  
GCACACTCCCACACTGAAACAAACAGTCATTCTATAGGACTTCCAGGAAGGATCTTCTTCT  
TAAGCTTCTTAATTATGGTACATCTCCAGTTGGCAGATGACTATGACTACTGACAGGAGAAT  
GAGGAAGTAGCTGGGAATATTCTGTTGACCAACATGGAGTCACCCATTCTTACTGGTA  
TTGGAAATAATAATTCTGAATTGCAAAGCAGGAGTTAGCGAAGATCTCATTCTTCCATG  
TTGGTGACAGCACAGTTCTGGCTATGAAAGTCTGCTTACAAGGAAGAGGATAAAAATCATAG  
GGATAATAATCTAAGTTGAAGACAATGAGGTTTAGCTGCATTGACATGAAGAAATTGA

GACCTCTACTGGATAGCTATGGTATTTACGTGTCTTTGCTTAGTTACTTATTGACCCAG  
CTGAGGTCAAGTATGAACTCAGGTCTCGGGCTACTGGCATGGATTGATTACATAACTG  
TAATTTAGCAGTGATTAGGGTTATGAGTACTTTGCAGTAAATCATAGGGTAGTAATG  
TTAATCTCAGGGAAAAAAAGCCAACCCTGACAGACATCCCAGCTCAGGTGGAAATC  
AAGGATCACAGCTCAGTGCAGGCTCCAGAGAACACAGGGACTCTCTTAGGACCTTATGT  
ACAGGGCCTCAAGATAACTGATGTTAGTCAGAAGACTTCATTCTGGCACAGTCAGCTG  
AGGCAATCCTGGAATTTCTCCGCTGCACAGTTCCAGTCATCCCAGTTGTACAGTTCTG  
GCACCTTTGGGTCAAGGCCGTGATCCAAGGAGCAGAAGTCCAGCTATGGTCAGGGAGTGCC  
TGACCGTCCAACTCAGTGCACCAAACAAAGGCGAAACCACAAGAGTGGCTTTGTTGAAA  
TTGCAGTGTGGCCCAGAGGGCTGCACAGTACTGGATTGACCACGAGGAAACATTAATCCT  
CAGCAAGTGCAATTGCAAGCATTAAATTGAACTAACGTACATAATGCAATCAGTATCA  
ACAAGTGGTTGGCTTGGAAAGATGGAGTCTAGGGCTCTACAGGAGTAGCTACTCTTAATG  
GAGTTGCATTTGAAGCAGGACACTGTGAAAAGCTGGCCTCTAAAGAGGCTGCTAAACATT  
AGGGTCAATTTCCAGTGCACTTCTGAAGTGTCTGCAGTTCCCATGCAAAGCTGCCAAA  
CATAGCACTCCAATTGAATAACAATTATGCAGGCGTACTGCTTCTGCCAGCAGTGCCT  
TCTCAAATGAACCAACAAACAATTCAAAGTCTAGTAGAAAGTAACAAGCTTGAATGTCA  
TTAAAAAGTATATCTGCTTTCAGTAGTCAGCTTATTTATGCCCACTAGAAACATCTGTAC  
AAGCTGAACACTGGGGCTCCAGATTAGTGGTAAAACCTACTTTATACAATCATAGAATCATA  
GAATGGCCTGGGTTGGAAAGGGACCCCAAGGATCATGAAGATCCAACACCCCCGCCAGGCA  
GGGCCACCAACCTCCAGATCTGGTACTAGACCAGGCAGGCCAGGGCTCCATCCAACCTGGCC  
ATGAACACCTCCAGGGATGGAGCATCCACAACCTCTGGCAGCCTGTGCCAGCACCTCAC  
CACCCCTCTGTGAAGAACCTTCCCTGACATCCAATCTAACGCTCCCTTGAGGTTAG  
ATCCACTCCCCCTTGTCTACTGTCTACTCTGTAAAAAGTTGATTCTCCTCCCTTTG  
GAAGGTTGCAATGAGGTCTCCTTGACGCCCTCTGCAGGATGAACAAAGCCCAGCT  
CCCTCAGCCTGTCTTATAGGAGAGGTGCTCCAGCCCTCTGATCATCTTGAGGCTCCCTC  
TGGACCCGCTCCAAGAGCTCCACATCTTCCCTGTACTGGGGGCCAGGCCTGAATGCAGTA  
CTCCAGATGGGCCTCAAAAGAGCAGAGTAAAGAGGGACAATCACCTCCTCACCTGCTGG  
CCAGCCCTCTGATGGAGCCCTGGATACAACACTGGCTTCTGAGCTGCAACTCTCCTTAT  
CAGTTCCACTATTAAAACAGGAACAATACAACAGGTGCTGATGGCCAGTCAGAGTTTCA  
CACTCTTCATTCGGTAGATCTTAGATGAGGAACGTTGAAGTTGCTCTGCGTGTGCTT  
CTTCCTCCTCAAATACTCCTGCCTGATACCTCACCCACCTGCCACTGAATGGCTCCATGGC  
CCCCTGCAGCCAGGGCCCTGATGAACCCGGCACTGCTTCAGATGCTGTTAATAGCACAGTA  
TGACCAAGTTGCACCTATGAATACACAAACAATGTGTTGCATCCTCAGCACTTGAGAAGAA  
GAGCCAAATTGCAATTGCTCAGGAAATGGTTAGTAATTGCCAATTAAAATTGTTATCT  
ACCATGGCTGTTTATGGCTGTTAGTAGTGGTACACTGATGATGAACAAATGGCTATGCAGT  
AAAATCAAGACTGTAGATATTGCAACAGACTATAAAATTCCCTGTGGCTTAGCCAATGTGG  
TACTTCCCACATTGTATAAGAAATTGGCAAGTTAGAGCAATGTTGAAGTGTGGAAAT  
TTCTGTATACTCAAGAGGGCTTTGACAACGTAGAACAGAGGAATCAAAAGGGGTGGG  
AGGAAGTTAAAAGAAGAGGCAGGTGCAAGAGAGCTTGCAGTCCCGCTGTGTACGACACTG  
GCAACATGAGGTCTTGCTAATCTGGTCTTGCCTGCCCTGGCTGCCTAGGGTGC  
GATCTGCCTCAGACCCACAGCCTGGCAGCAGGAGGACCTGATGCTGCTGGCTCAGATGAG  
GAGAATCAGCCTGTTAGCTGCCTGAAGGATAGGCACGATTGGCTTCTCAAGAGGAGT  
TTGGCAACCAGTTTCAGAAGGCTGAGGACCATCCCTGTGCTGCACGAGATGATCCAGCAGATC  
TTAACCTGTTAGCACCAAGGATAGCAGCGCTGCTGGATGAGGACCTGCTGGATAAGTT  
TTACACCGAGCTGTACCGAGCTGAACGATCTGGAGGCTTGCAGGAGTCCAGGGCGTGGCG  
TGACCGAGACCCCTCTGATGAAGGAGGATAGCATCCTGGCTGTGAGGAAGTACTTCAAGAGG  
ATCACCCCTGTACCTGAAGGAGAAGAAGTACAGCCCCCTGCGCTTGGGAAGTCGTGAGGGCTGA  
GATCATGAGGAGCTTAGCCTGAGCACCAACCTGCAAGAGAGCTTGGAGGTCTAAGGAGTAAA  
AAGTCTAGAGTCGGGGCGGCCGCTTCGAGCAGACATGATAAGATAACATTGATGAGTT  
GGACAAACCACAACTAGAATGCAGTGAAGGGGGAGGTGTGGGAGGTTTAAAGCAAGTAAA  
ACCTCTACAAA  
TTATGTTAGGTTCAAGGGGGAGGTGTGGGAGGTTTAAAGCAAGTAAA  
ACCTCTACAAA

TGTGGTAAAATCGATAAGGATCCGTCGACCGATGCCCTGAGAGCCTCAACCCAGTCAGCT  
CCTTCCGGTGGCGCGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTATCATG  
CAAACCGTAGGACAGGTGCCGGCAGCGCTCTCCGCTCGCTCACTGACTCGCTGCGCT  
CGGTCGTTGGCTCGCGGAGCGGTATCAGCTCACTCAAAGGCGGTAAACGGTTATCCACA  
GAATCAGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG  
TAAAAAGGCCGCGTTGGCGTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAA  
ATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACAGGCCTTCCC  
CCTGGAAGCTCCCTCGTGCCTCTCCGTTCCGACCCCTGCCGCTACCGGATACTGTCCGC  
CTTCTCCCTCGGAAAGCGTGGCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTGG  
TGTAGGTCGTTCGCTCCAAGCTGGCTGTGCACGAACCCCCGTTAGCCGACCGCTGC  
GCCTTATCCGTAACATCGTCTTGAGTCCAACCCGTAAGACACGACTATGCCACTGGC  
AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCAGGTACAGAGTTCTGA  
AGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTTGCTGAAG  
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGAAACAAACCACCGCTGGTAG  
CGGTGGTTTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTAAGAAGATC  
CTTGATCTTCTACGGGTCTGACGCTCAGTGGAACGAAAACACGTTAAGGGATTTG  
GTCATGAGATTATCAAAAAGGATCTCACCTAGATCCTTAAATTAAAAATGAAGTTTAA  
ATCAATCTAAAGTATATGAGTAAACTGGCTTGACAGTTACCAATGCTTAATCAGTGAGG  
CACCTATCTCAGCGATCTGTCTATTGTTCATCCATAGTTGCCTGACTCCCCGTGTAG  
ATAACTACGATAACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAACCGCAGACCC  
ACGCTCACCGGCTCCAGATTATCAGCAATAAACAGCCAGCCGGAAAGGGCCAGCGCAGAA  
GTGGTCCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGTTGCCGGAAAGCTAGAGTA  
AGTAGTTGCCAGTTAATAGTTGCGAACGTTGCTACAGGCATCGTGGTAG  
ACGCTCGTCTGGTATGGCTTCATTAGCTCCGGTCCAAACGATCAAGGCGAGTTACAT  
GATCCCCCATGTTGCAAAAAAGCGTTAGCTCCTCGGTCCATCGTGGTAG  
AAGTTGGCCGCAGTGTATCACTCATGGTTATGGCAGCAGTCATAATTCTTACTGTCAT  
GCCATCCGTAAGATGCTTCTGTGACTGGTAGTACTCAACCAAGTCATTGAGAATAGT  
GTATGCGGCGACCGAGTTGCTCTGCCGGCGTCAATACGGATAATACCGGCCACATAGC  
AGAACTTAAAAGTGTCTCATATTGAAAACGTTCTCGGGCGAAAACCTCTCAAGGATCTT  
ACCGCTGTTGAGATCCAGTTGATGTAACCCACTCGTGCACCCAACTGATCTCAGCATCTT  
TTACTTCAACCAGCGTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCCGAAAAAGGGA  
ATAAGGGCGACACGGAAATGTTGAATACTCATACTCTCCTTTCAATATTATTGAAGGCAT  
TTATCAGGGTTATTGTCATGAGCGGATAACATATTGAATGTATTAGAAAAATAACAAA  
TAGGGGTTCCGCGCACATTCCCCGAAAAGTGCACCTGACGCCCTGTAGGGCGCATTA  
AGCGCGGCGGGTGTGGTTACGCGCAGCGTACCGCTACACTGCCAGGCCCTAGCGCC  
CGCTCCTTCGCTTCTCCCTCCTTCGACCGTACCGCTACACTGCCAGGCCCTAGCGCC  
TAAATCGGGGGCTCCCTTAGGGTTCCGATTAGTGTCTTACGGCACCTCGACCCAAAAAA  
CTTGATTAGGGTGATGGTCACGTAGTGGCATCGCCCTGATAGACGGTTTCGCCCTT  
GACGTTGGAGTCCACGTTCTTAATAGGACTCTGTTCAAACGGAAACAACTCAACC  
CTATCTCGGTCTATTCTTTGATTATAAGGATTTCGCCATTGCGCTATTGGTTAAAA  
AATGAGCTGATTTAACAAAATTAAACGCGAATTAAACAAATATTACGTTACAATTTC  
CCATTGCCATTAGGCTCGCAACTGTTGGGAAGGGCGATCGGTGCGGCCCTTCGCTAT  
TACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAAGGTACGGGAGGTACTGGAGCG  
GCCGCTCTAGAACTAGGATCCCCGGCGAATAAAATATCTTATTTCATTACATCTG  
TGTGTTGGTTTTGTGTGAATCGATAGTACTAACATACGCTCTCCATCAAACAAACGAA  
ACAAAACAAACTAGCAAAATAGGCTGCCCCAGTGCAAGTGCAGGTGCCAGAACATTCTCT  
ATCGATAGGTACCGAGCTTACGCGTCTAGCCCTCGAGCAGGATCTACATTGAATCAA  
TATTGGCAATTAGCCATTAGTCATTGGTTATAGCATAAAATCAATTGGCTATTGGCC  
ATTGCATACGTTGTATCTATACATAATATGTACATTATGGCTCATGTCCAATATGAC  
GCCCATGTTGACATTGATTAGGACTAGTTATTAGTAATAGTAATCAATTACGGGGTCATTAGTT  
CATAGCCCATTGGAGTTCCGCGTTACATAACTACGGTAAATGGCCCGCTGGCTGACC  
GCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCCATTAGTAACGCCAATAG

GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCACTGGCAGTACAT  
CAAGTGTATCATATGCCAAGTCCGCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTG  
GCATTATGCCAGTACATGACCTTACGGGACTTCCTACTGGCAGTACATCTACGTATTAG  
TCATCGCTATTACCATGGTATGCGGTTGGCAGTACATCAATGGCGTGGATAGCGGTTT  
GAECTCACGGGATTCCAAGTCTCCACCCATTGACGTCAATGGAGTTGGCACCA  
AAATCAACGGACTTCAAAATGTCGAACAACCTCCGCCATTGACGCAAATGGCGGTA  
GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTAGTGAACCGTCAGATCGCCTGG  
AGACGCCATCCACGCTGTTGACCTCCATAGAACACCCGGGACCGATCCAGCCTCCCTC  
GAAGCTCGACTCTAGGGCTCGAGATCTGCATCTAAGTAAGCTTGCATGCCTGCAGGTCGG  
CCGCCACGACCGGTGCCACCACATCCCCTGACCCACGCCCTGACCCCTCACAGGAGACG  
ACCTTCCATGACCGAGTACAAGCCCACGGTGCCTCGCCACCCGCGACGACGTCCCCCGGG  
CCGTACGCACCCTGCCGCCCGTTCGCCGACTACCCGCCACGCCACACCGTCACCCG  
GACCGCCACATCGAGCGGGTACCGAGCTGCAAGAACTCTCCTCACGCCGTCGGCTCGA  
CATCGGCAAGGTGTGGTGCACGGACGGCGCGGTGGCGGTCTGGACCACGCCGAGA  
GCGTCAAGCGGGCGGTGTCGCCGAGATCGGCCGCGATGGCCGAGTTGAGCGGTTCC  
CGGCTGGCCGCGCAGAACAGATGGAAGGCCCTGGCGCCGACCGGCCAAGGAGCCGC  
GTGGTTCTGGCCACCGTCGGCGTCTCGCCGACCACCGAGGCAAGGGCTGGCAGCGCCG  
TCGTGCTCCCCGGAGTGGAGGCGGCCAGCGCGCCGGTGCCTGGAGACCTCC  
GCGCCCCGCAACCTCCCCCTACGAGCGGCTCGGCTTACCGTCACCGCCGACGTCGAGGT  
GCCCGAAGGACCGCGCACCTGGTGCATGACCGCAAGCCGGTGCCTGACGCCGCCCCACG  
ACCCGCAGCGCCGACCGAAAGGAGCGCACGACCCATGGCTCCGACCGAAGCCGACCCGGG  
CGGCCCCGCCGACCCGCACCCGCCCGAGGCCACCGACTCTAGAGTCGGGCGGCCG  
CGCTTCGAGCAGACATGATAAGATACTTGATGAGTTGGACAAACCACAAACTAGAATGCAG  
TGAAAAAAATGCTTATTGTGAAATTGTGATGCTATTGCTTATTGTAACCATTATAAG  
CTGCAATAACAAGTTACAACAACAATTGCATTCTATTGTTACGGTCAAGGTTAGGGGAGG  
TGTGGGAGGTTTTAAAGCAAGTAAACCTCTACAAATGTGGTAAATCGATAAGGATCAA  
TTCGGCTTCAGGTACCGTCACGATGTAGGTACGGTCTCGAACCGCGGTGCCAG  
GGCGTGCCTGGGCTCCCCGGCGTACTCCACCTCACCCATCTGGTCCATCATGATGAA  
CGGGTCGAGGTGGCGGTAGTTGATCCCGCGAACCGCGCCGGAAAGCCCTGCCCT  
CGAAACCGCTGGCGCGGTGGTCACGGTGAGCACGGACGTGCGACGGCGTCGGCGGGTGC  
GATAACGGGGCAGCGTCAGCGGTTCTCGACGGTACGGGGCATGTCGACAGCCGAATT  
GATCCGTGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTCCGGTGGCGCGGG  
GCATGACTATCGTCGCCGACTTATGACTGCTTCTTATCATGCAACTCGTAGGACAGGTG  
CCGGCAGCGCTTCCGCTCGCTCACTGACTCGCTCGCTCGGTGGCTGGCTGCGGC  
GAGCGGTATCAGCTCACTCAAAGCGGTAATACGGTTATCCACAGAATCAGGGATAACGCA  
GGAAAGAACATG

**Fig. 15**

**pOM IFN-Ins-CMV-pur-attB (SEQ ID NO: 8)**

GGCCGCCACCGCGGTGGAGCTCCAATTGCCCTATAGTGAGTCGTATTACAATTCACTGGCC  
GTCGTTTACAACGTCGTGACTGGAAAACCCCTGGCGTTACCCAACTTAATGCCCTGCAGC  
ACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCACCGATGCCCTCCAAAC  
AGTTGCGCAGCCTGAATGGCGAATGGGACGCCCTGTAGCGGCATTAAGCGCGGGGGT  
GTGGTGGTTACGCGCAGCGTACACTGCCAGCGCCCTAGCGCCGCTCCCTTCGC  
TTCTCCCTTCTCGCCACGTTGCCGGCTTCCCCGTCAAGCTCTAAATCGGGGGC  
TCCCCTTAGGGTCCGATTAGTGTCTTACGGCACCTCGACCCAAAAACTGATTAGGGT  
GATGGTTCACGTAGTGGGCCATGCCCTGATAGACGGTTTCGCCCTTGACGTTGGAGTC  
CACGTTCTTAATAGTGACTCTGTTCAAACCTGGAACAAACACTCAACCCATTCTCGGTCT  
ATTCTTTGATTATAAGGGATTTGCCGATTCGCCCTATTGGTTAAAAAATGAGCTGATT  
TAACAAAAATTAAACGCGAATTAAACAAATATTAAACGCTTACAATTAGTGGCACCTTT  
CGGGGAAATGTGCGCGAACCCCTATTGTTATTCTAAATACATTCAAATATGTATCC  
GCTCATGAGACAATAACCTGATAAATGCTCAATAATATTGAAAAAGGAAGAGTATGAGTA  
TTCAACATTCCGTGCGCCCTTATTCCCTTTGCGGCATTTGCCTCCTGTTGCT  
CACCCAGAAACGCTGGTAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTA  
CATCGAACTGGATCTCAACAGCGGTAAGATCCTGAGAGTTTCGCCCGAAGAACGTTTC  
CAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGCGTATTATCCGTATTGACGCCGGG  
CAAGAGCAACTCGTCGCCGCATAACTATTCTCAGAATGACTTGGGTGAGTACTCACCAGT  
CACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTATGCAGTGCTGCCATAACCA  
TGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAAC  
GCTTTTTGCACAACATGGGGATCATGTAACCGCCCTGATCGTTGGAACCGGAGCTGAA  
TGAAGCCATACCAAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACACGTTGC  
GCAAACATTAACTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAAATAGACTGGATG  
GAGGGGATAAAGTTGCAGGACCACTCTCGCTCGGCCCTCCGGCTGGCTGGTTATTGC  
TGATAAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGAGCAGTACTGGGCCAGATG  
GTAAGCCCTCCGTATCGTAGTTATCTACACGACGGGAGTCAGGCAACTATGGATGAAACGA  
AATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAAGCATTGGTAACTGTCAGACCAAGT  
TTACTCATATATACTTAGATTGATTAAAACCTTCAATTAAATTAAAAGGATCTAGGTGA  
AGATCCTTTGATAATCTCATGACCAAAATCCCTAACGTGAGTTTCGTTCACTGAGCG  
TCAGACCCCGTAGAAAAGATCAAAGGATCTCTGAGATCCTTTCTGCGCGTAATCTG  
CTGCTTGCAAACAAAAACCAACCGCTACCGCGTGGTTGTTGCCGATCAAGAGCTAC  
CAACTCTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATAACAAATCTGCTCTCTA  
GTGTAGCCGTAGTTAGGCCACCACTCAAGAACTCTGTAGCACCCTACATACCTCGCTCT  
GCTAATCCTGTTACCACTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTGGACT  
CAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGCTGAACGGGGTTCGTGCACACAG  
CCCAGCTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG  
CGCCACGCTCCCGAAGGGAGAAAGGGGACAGGTATCCGTAAGCGGCAGGGTGGAAACAG  
GAGAGCGACGAGGGAGCTCCAGGGGAAACGCCCTGGTATCTTATAGTCCTGTCGGGTT  
GCCACCTCTGACTTGAGCGTCGATTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAA  
AAACGCCAGCAACCGGGCTTTACGGTTCTGGCCTTTGCTGGCTTGCTCACATGT  
TCTTCTGCGTTATCCCTGATTCTGAGATAACCGTATTACCGCCTTGAGTGAGCTGAT  
ACCGCTCGCCGAGCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCG  
CCCAATACGCAAACCGCCTCTCCCCCGCGTTGGCCGATTCAATTAAATGCAGCTGGCACGACA  
GGTTCCCGACTGGAAAGCGGGCAGTGAGCGAACGCAATTAAATGTGAGTTAGCTCACTCAT  
TAGGCACCCAGGCTTACACTTATGCTTCCGGCTCGTATGTTGTGGAATTGTGAGCGG  
ATAACAATTACACAGGAAACAGCTATGACCATGATTACGCCAGCTCGAAATTAAACCTC  
ACTAAAGGGAACAAAAGCTGGGTACCGGGCCCCCTCGACTAGAGGGACAGCCCCCCCCCA  
AAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGGCAGCAGCGAGGCCGGGG  
CTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCAGGGGACAGCCC  
GCACGGGAAAGGTGGCACGGATCGCTTCTGAACGCTCTCGCTGCTTTGAGCCTG

CAGACACCTGGGGGATA CGGGAAAAAGCTT TAGGCTGAAAGAGAGATTAGAATGACAGA  
ATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTAT  
GTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATG  
CCTGCAGGGATGGGGCATCCACAGCCTCCTGGCAACCTGTTAGTGCAGTCACCACCCCT  
GGGGAAAAACTGCCTCCTCATATCCAACCAAACCTCCCCTGTCTCAGTGTAAAGCCATT  
CCCCTGTCCTATCAAGGGGAGTTGCTGACATTGTTGGTCTGGGTGACACATGTTG  
CCAATTCAAGTGCATCACGGAGAGGCAGATCTTGGGATAAGGAAGTGCAGGACAGCATGGAC  
GTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTAGAACAGCC  
TTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAG  
CACAAAAAGGCCACAGACACTGCTGGCCCTGTGCTGAGCCTGCATGTTGATGGTGTCTG  
GATGCAAGCAGAAGGGTGGAAAGAGCTTGCCTGGAGAGATAACAGCTGGGTAGTAGGACTGG  
GACAGGCAGCTGGAGAATTGCCATGTAGATGTTCATACAAATCGTCAAATCATGAAGGCTGGA  
AAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCGTGCCTGACCCCTCCATGTCC  
CTCAGTGCCACATCCCCACAGTTCTCATCACCTCCAGGGACGGTGACCCCCCCTCCGT  
GGGCAGCTGTGCCACTGCAGCACCGCTCTTGGAGAAGGTAAATCTGCTAAATCCAGCCCG  
ACCCCTCCCTGGCACACGTAAAGGCCATTATCTCATCCAACCTCCAGGACGGAGTCAGTGA  
GGATGGGGCTCTAGTCGAGGTGACGGTATCGATAAGCTTGTAGGAGCAATAGGACT  
CTCAACCTCGTGAGTATGGCAGCATGTTACTCTGCACTGGAGTCCAGCGTGGAAACAATC  
TGCCTTGACATGAGTCTCGTGGCCAATATTCCCCAACGGTTTCCTCAGCTGTCTTG  
TCTCCTAAGCTCTAAAACACCTTTGGTAATAAACTCACTGGCAACGTTATCTGTCT  
TACCTTAGTGTACGTTCATCCCTATTCCCCCTTCCTCCGTGTGGTACACAGTGGT  
GCACACTGTTCTCTGTTGATGTTCTGCTCTGACAGCCAATGTGGTAAAGTTCTCCTGC  
CACGTGTGTGTTGTTCACTCAAAAAGGGCCCTGGGCTCCCTGGAGCTCTCAGGCA  
TTCCCTTAATCATCACAGTCACGCTGGCAGGATTAGTCCCTCCTAAACCTTACAATGACCTG  
AACGTGTGCTCCCTTTGTTAGTCAGTGCAGGGAGACGTTGCCTCAAGATCAGGGTCCATC  
TCACCCACAGGGCATTCCAAGATGAGGTGGATGGTTACTCTCACAAAAAGTTTCTTAT  
GTTGGCTAGAAAGGAGAACTCACTGCCTACGTGAATTCCCTAGTCCTGGTTCTGCTGC  
CACTGCTGCCTGTGCAGCCTGCCCAGGATGGAGGGCAGCAACTGCTGTACAAAGGTGATCC  
CACCTGTCTCCACTGAAATGACCTCAGTGCACGTGTTGATAGGGTATAAGTACGGGAG  
GGGATGCCCGGCTCCCTCAGGGTTGCAGAGCAGAAGTGTCTGTGATAGAGTGTCTTA  
ATCTATTAATGTAACAGAACAACTTCAGTCCTAGTGTGTTGTGGCTGGAATTGCCATGTG  
GTAGGGACAGGCCTGCTAAATCACTGCAATCGCCTATGTTCTGAAGGTATTGGAAAGAAA  
GGGATTGGGGATTGCCTGTGATTGGCTTAATTGAATGGCAAATCACAGGAAAGCAGTTC  
TGCTCAACAGTTGGTTGTTCAGCCATTCTGCAGCCAAAGAGGCCGGGTGCCAGCGATAT  
AATAGTTGTCACTGTGCTGTATGGATGACAGGGAGGTAGGGTACCTGAGGACCACCC  
CAGCTTCTGCTAGCGTAGGTACAGTCACCACCTCCAGCTCACACGAGTCCCACGTGGTT  
ACCAAAGAAACACAATTATTGGACCAGTTGGAAAGTCACCCGCTGAATTGTGAGGCTAGA  
TTAATAGAGCTGAAGAGCAAATGTTCCAACCTGGAGATACTAGTTGGTATTAGTATCAGAG  
GAACAGGGCCATAGCACCTCCATGCTATTAGATTCCGGCTGGCATGTACTTTCAAGATGAT  
TTGTAACATAACATGGCTTATTGTGCTTGAAGTCTGTGCTTAATGTAATGTTCTT  
TGGTTTATATAACCTTCTGCCATTGCTCTCAGGTGTTCTGCAGAACACTGGCTGCTTT  
AATCTAGTTAACTGTTGCTTGATTATTCTTAGGGATAAGATCTGAATAAAACTTTGTGGC  
TTTGGCAGACTTAGCTGGCCTAGCTCCACATTAGCTTTGCTGCCTTGTGAAGC  
TATCAAGATCCTACTCAATGACATTAGCTGGGTGCAGGTGACCAAATCTGCTCTGTGGAA  
CACATTGTCTGATGATACCGAACAGTGAACCTCAAAGAGGCACAGAGTTAGAAGAAG  
TCTGTGCAATTCAAGAGAAAAGCCAAAGTGGCATTAGACACACTTCCATGCAGCATTG  
CAGTAGGTTCATATAAAACTACAAAATGGAATAAACCAACTACAAATGGAAAAGCCTGATA  
CTAGAATTAAATATTCAACCAGGCTCAAGGGGTGTTCATGGAGTAATATCACTCTATAAA  
AGTAGGGCAGCCAATTATTCAACAGACAAAGCTTTTTCTGTGCTGCAGTGTGTTTT  
CGGCTGATCCAGGGTTACTTATTGTGGGTCTGAGAGCTGAATGATTCTCCTGTGTCATGT  
TGGTGAAGGAGATATGCCAGGGGAGATGAGCATGTTCAAGAGGAAACGTTGCATTGGT  
GGCTGGGAGAAAGGTAGAACGATATCAGGTCCATAGTGTCACTAAGAGATCTGAAGGATGG

TTTTACAGAACAGTTGACTTGGCTGGGTGCAGGCTGGCTGTAAATGGATGGAAGGATGGAC  
AGATGGTGGACAGAGATTCCTGTGCAGGAGATCATCTCCTGAGCTCGGTGCTTGACAGACT  
GCAGATCCATCCCATAACCTCTCCAGCATGAGAGCGCGGGAGCTTGGTACTGTTAGTCAGTC  
TGCTGCTTGTGCTCCTGGGTGCACAGTGGTGATTTCTTACTCACACAGGGCAAAACCT  
GAGCAGCTCAAAGTGAACAGGTTGCTCTCATAGGCCATTAGTTGTCAAGATGAGGTTTT  
GGTTCTTGTGTTGTAAGGTGGGAAGAAGCACTGAAGGATCAGTTGCGAGGGCAGGGTTA  
GCACTGTTAGAGAAGTCTTAACTCCTCTCATGAACAAAAAGAGATGCAGGTGAGA  
TTCTGGCAAGCATGCAGTGAAGGAGAAAGCCCTGAATTGATATATGTGCAATGTTGGC  
ACCTAACATTCCCCGCTGAAGCACAGCAGCTCCAGCTCCATGCAGTACTCACAGCTGGTGCA  
GCCCTCGGCTCCAGGGTCTGAGCAGTGCTGGACTCACGAGGTTCCATGTCTTCACACTGA  
TAATGGTCCAATTCTGGAATGGGTGCCCATCCTGGAGGTCCCCAAGGCCAGGCTGGCTGC  
GTCTCCGAGCAGCCGATCTGGTGGTAGCCAGCCATGGCAGGAGTTAGAGCCTGATG  
GTCTTTAAGGTCCCTCCAACCTAACGCCATCCTACGATTCTAGGAATCATGACTTGTGAGTG  
TGTATTGCAGAGGAATATTAAAGTTATAAATGTTCTCCCTCCTGTTGTCAAAG  
TTATCTTGATCGCCTTATCAATGCTTGGAGTCTCCAGTCATTCTTACAMCAAAAGA  
GGAGGAAGAATGAAGAGAATCATTAATTCTGATTGAATAGTAGGATTAGAAAGCTGTA  
CGTAATGCCGTCTTTGTATCGAGCTGTAAAGGTTCTCATCATTATCAGCGTGGTACATA  
TCAGCACTTTCCATCTGATGTGGAAAAAAATCCTTATCATCTACAGTCTGTACCTAA  
ACATCGCTCAGACTCTTACCAAAAAAGCTATAGGTTAAAACATACATCTGCTGATAATT  
GCCTGTTTAGCTCTTCCATATGCTGCGTTGTGAGAGGTGCGTGGATGGGCCTAAAC  
TCTCAGCTGCTGAGCTGATGGTGCTTAAGAATGAAGCACTCACTGCTGAAACTGTTCA  
TTTCACAGGAATGTTAGTGGCATTGTTTATAACTACATATTCTCAGATAAAATGAAAT  
CCAGAAATAATTGCAAACACTGCATCCGTTGCACAGGTCTTATCTGCTAGCAAAGGA  
AATAATTGGGGATGGAAAAACATTCCCTCAGACATCTATATTAAAGGAATATAATCCTG  
GTACCCACCCACTCATCCCTCATTATGTCACACTCAGAGATACTCATTCTCTGTTTA  
TCATTTGATAGCGTTTCTTGTTCTTGCCACGCTCTGGCTATGGCTGCACGCTCTGCA  
CTGATCAGCAAGTAGATGCGAGGGAAGCAGCAGTGAGAGAGGGCTGCCCTCAGCTGGCACCC  
GCCGCTCAGCCTAGGAGGGACCTTGCCTTCCACCAGCTGAGGTGCAGCCCTACAAGCTTA  
CACGTGCTGCGAGCAGGTGAGCAAAGGGAGTCTTACGGTGTGTTCTGCTGCCGGAAAGC  
AAAACTTACTTCATTCCCTTGAAGAATGAGGAATGTTGGAAACGGACTGCTTTA  
CGTTCAATTCTCTTCCCTTAAGGCTCAGCCAGGGCATTGCTGAGGACGGCATCGGG  
GCCCTGGACCAATCTGTCCTGAAGGAAGGAACGTGCCTTCCAAGTGCCAGCCCCACAGCCCC  
AGCCCTCCCTGTGCTCCAATTCTCCTCTCCCTCTCCCTTGCTGTTGTG  
TCGGGTAGAAATCATGAAGATTAGAAGAGAAAACAAAATACTGGAGTGGAAACCCAGGTG  
ATGCAGTTCATTCACTGTCAGCTGTCAGGTTGCGTGTGCTATAGGTCTGTATCAGAGATGCTARC  
ACCACTTGCTGTCGGTGCTTAACTCGGGTGAACCTCCTTCACTCGCATTGCGCTCTGCA  
TTATTACATCCCAGCATCCATTCCCTCTGGAAAATGGCGCACTGGATCTTAATGGA  
AGACTTCCCTTTCAAGGCCTGTGGATGTGCAGTGACAAGAAACGTGGAGGGCTGAGC  
AGCAGCACTGCCCAAGGGAGCAGGAGCAGGAGCGGATGCCATGGCAGCATCCAAATGATGTC  
AGCGGATGCTGAGCAGGCAGCGACGAACGGACAGAAGCGATGCGTACACCTCTGTTGACA  
TGGTATTGGCAGCGATTAAACACTCGCTTAGTCCTGCTATTCTCACAGGCTGCATT  
AAATGAACGAAGGGAGGGAAAAAGATGCAAAATCCGAGACAAGCAGCAGAAATATT  
CTTCGCTACGGAAAGCGTGCCTAACACCTCTCCAAACAGCACCGAGAGCACAGCGTAAC  
CTTTTCAAGACCAGAAAAGGAAATTACAAAGCCTCTGGATACCAGCGCGTTCAGCTCT  
CCTGATAGCAGATTCTGTCAGGTTGCGAATGGGTATGGTGCAGGAGGTGCAGGGACCA  
TATGATCATATACAGCACAGCAGTCATTGTCATGTATTAAATATATTGAGTAGCAGTGTT  
ACTTTGCCAAAGCAATAGTTCAAGAGATGAGTCCTGCTGCATACCTCTATCTTAAACTAACT  
TATAAATAGTAAAACCTCTCAGTCAGCCACGTGCTCCTCTGTCAAGCACCAATGGTGCT  
TCGCCCTGCACCCAGCTGCAAGGAATCAGCCCGTGAATCTCATTAAACACTCAGCTCTGCAAGGAT  
AAATTAGATTGTTCCACTCTTTGTTAATTACGACGGAACAATTGTTAGTGCTGAT  
GGCCTTAATTGTCAGCTACAGAAAACGTCTCCATGCAGTTCTGCGCCAGCAAACGTGTC

CAGGCTATAGCACCGTGATGCATGCTACCTCTCACTCCATCCTCTCTTTCCCACCAGG  
GAGAGCTGTGTGTTTCACTCTCAGCCACTCTGAACAAATACCAAACGCTACGCAC TGCC  
CCTCGGAAAGAGAATCCCCTGTTGCTTTTATTACAGGATCCTCTTAAAGCAGACC  
ATCATTCACTGCACAAACCCAGAGCTCATGCCTCTCCTCCACAACCGAAACAGCCGGCTC  
ATTTGTCTTTAAATGCTGTTCCAGGTGAATTGGCCAGCGTGTGGCTGAGATCCA  
GGAGCACGTGTCAGCTTCTGCTCATTGCTCCTGTCATTGCCTCTTCTGGGGTT  
CCAAGAGGGGGGAGACTTGCGCGGGATGAGATAATGCCCTTCTAGGGTGGCTGCT  
GGCAGCAGAGTGGCTCTGGTCACTGTGGCACCAATGGGAGGCACCAGTGGGGTGTGTT  
TGTGCAGGGGGGAAGCATTACAGAATGGGGTGATCCTGAAGCTGCACTCCAAGGCTTG  
TCTGTGTACCCAGTGAAATCCTCCTGTTACATAAGCCCAGATAGGACTCAGAAATGTA  
GTCATTCCAGCCCCCTCTCCTCAGATCTGGAGCAGCACTGTTGCAGCCAGTCCTCCCC  
AAAATGCACAGACCTCGCCGAGTGGAGGGATGTAAACAGCGAAGGTTAATTACCTCCTG  
TCAAAAACACTTGTGGTCCATAGATGTTCTGTCAATCTACAAAACAGAACCGAGAGGCA  
GCGAGCACTGAAGAGCGTGTCCCAGCTGAGTTAATGAGACTTGGCAGCTCGTGTGCAGA  
GATGATCCCTGTGCTTCATGGGAGGCTGTAACCTGTCTCCCCATCGCCTCACACCGCAGTG  
CTGTCCTGGACACCTCACCCCTCATAAGCTGTAGGATGCAGCTGCCAGGGATCAAGAGACT  
TTTCCTAAGGCTCTAGGACTCATCTTGCCGCTCAGTAGCGTGCAGCAATTACTCATCCCA  
ACTATACTGAATGGGTTCTGCCAGCTGCTGTTGTCAATAAGCATTCTTCATTTGC  
CTCTAAGTTCTCAGCAGCACCGCTCTGGGTGACCTGAGTGGCACCTGGAACCCGAGGG  
GCACAGCCACCACCTCCCTGCTGCTGCCAGGGACTCATGTGCTGCTGGATGGGGGA  
AGCATGAAGTTCCACCCAGACACCTGGGTGCAATGGCTGCAGCGTGCCTTCTGGTAT  
GCAGATTGTTCCAGCCATTACTGTAGAAATGTGCTGTGGAAGCCCTTGTATCTCTTCT  
GTGCCCTTCAGCAAAAGCTGTGGAAAGCTCTGAGGCTGCTTCTGGGTGAGGAAT  
TGTATGTTCTTAAACAAAATTATCCTTAGGAGAGAGCACTGTGCAAGCATTGTGCAC  
ATAAAACAATTCAAGGTTGAAAGGGCTCTGAGGTTCCAGCCTGACTACTGCTCGAAGCA  
AGGCCAGGTTCAAAGATGGCTCAGGATGCTGTGCCTCCTGATTATCTGTGCCACCAATG  
GAGGAGATTCACAGCCACTCTGCTTCCGTGCCACTCATGGAGAGGAATATTCCCTTATATT  
CAGATAGAATGTTATCCTTAGCTCAGCCTCCCTATAACCCATGAGGGAGCTGCAGATCC  
CCATACTCTCCCTCTGGGGTGAAGGCCGTGCCCCAGCCCCCTCCACCTGTGC  
CCTAACGCAGCCGCTGGCCTCTGCTGGATGTGTCCTATATGTCAATGCCTGTCTGCAGT  
CCAGCCTGGACATTAAATTCAACCAGGGTAATGTGGAACTGTGTCATCTCCCTGCAG  
GGTACAAAGTTCTGCACGGGCTCTCGGTTAGGAAAACCTTCACTGGTGTACCTGAAT  
CAAGCTCTATTAAAGTTCATAGCACATGGATGTGTTCTAGAGATACTGTTAAATG  
GTATCAGTGTGTTTATTGCTTGTGCTTACTTCAAACAGTCGCTTGGCAGGAGGTGA  
GGGACGGGCTGCCAGAAGAACAGATGGAAATTGCAGATTGAGTCACGTTAACAGCAG  
GCATCTGGAGTGATTGAGGCAGTTCATGAAAGAGCTACGACCACCTATTGTTGTTCC  
CCTTTACAACAGAAAGTTCATCAAAATAACGTGGCAAAGCCCAGGAATGTTGGAAAAG  
TGTAGTTAAATGTTGTAATTCAATTGCTGGAGTGCTACCAAGCTAAGAAAAAGTCCTACC  
TTTGGTATGGTAGTCCTGCAGAGAATACAACATCAATATTAGTTGGAAAAAAACACCACCA  
CCACCAGAAACTGTAATGGAAATGAAACCAAGAAATTCCCTGGTAAGAGAGAACAGGATG  
TCGTATACTGGCCAAGTCCTGCCAGCTGTCAAGCCTGCTGACCCCTCTGCAGTTCA  
GAAACGTGGCACTGTAAGACGTGTCCTGCTGCCACAGATCTCTGCCCTGTG  
CTGACTCCTGCACACAAGAGCATTCCCTGTAGCCAAACAGCGATTAGCCATAAGCTGCACC  
TGACTTGAGGATTAAGAGTTGCAATTAAAGTGATTGAGTCAGCAGGAGATCAGTGGCAGGGTT  
GCAGATGAAATCCTTCTAGGGTAGCTAAGGGCTGAGCAACCTGCTACAGCACAAGCC  
AAACCAGCCAAGGGTTCTGCTGTTCACAGAGGCAGGGCCAGCTGGAGCTGGAGGG  
TTGTGCTGGGACCCCTCTCCCTGTGCTGAGAATGGAGTGATTCTGGGTGCTGTTCTGTG  
CTTGCAGCTGAGCAGCTCAAGGGAGATCGGTGCTCATGCAGTGCCAAAACCTCGTGTG  
TGCAGAAAGATGGATGTGCACCTCCCTGCTAATGCAGCCGTGAGCTATGAAGGCAATG  
AGCCCTCAGTGCAGCAGGAGCTGTAGTGCACCTGTAGGTGCTAGGGAAAATCTCTGGTTC  
CCAGGGATGCATTCAAGGGCAATATATCTTGAGGCTGCCAAATCTTCTGAAATATTC



GTGCGCCTGCCACCCGCGACGACGTCCCCGGGCCGTACGCACCCCTGCCGCCGCGTCGC  
CGACTACCCGCCACGCCACACCGTCGACCCGGACCGCACATCGAGCGGGTCACCGAGC  
TGCAAGAACTCTTCCTCACGCCGTGGGCTCGACATCGCAAGGTGTGGCTGCCGGACGAC  
GGCGCCGCCGGTGGCGGTCTGGACCACGCCGGAGAGCGCTGAAGCGGGGGCGGTGTTGCCGA  
GATCGGCCCGCGCATGGCGAGTTGAGCGGTCTCCGGCTGCCGCCAGCAACAGATGGAAG  
GCCTCCTGGCGCCGACCGGCCAAGGAGCCCGTGGTCTGGCCACCGTCGGCGTCTCG  
CCCGACCACCAGGGCAAGGGCTGGGCAGCGCCGTGCTCCCCGGAGTGGAGGGCGGCCGA  
GCGGCCGGGTGCCGCCCTCTGGAGACCTCCGCCCGCAACCTCCCCTTACGAGC  
GGCTCGGCTCACCGTCACCGCCGACGTGAGGTGCCGAAGGACCGCGCACCTGGTGCATG  
ACCCGCAAGCCCAGGTGCCTGACGCCGCCACGACCCGCAGCGCCGACCGAAAGGAGCGC  
ACGACCCCATTGGCTCCGACCGAACGCCGACCCGGCGGCCGACCCGCACCCGCC  
GAGGCCACCAGACTCTAGAGTCGGGGCGGCCGCTCGAGCAGACATGATAAGATAACAT  
TGATGAGTTGGACAAACCACAACCTAGAATGCAGTGAAAAAAATGCTTATTGTGAAATT  
GTGATGCTATTGCTTATTGTAACCATTATAAGCTGCAATAAACAGTAACAACAACAA  
TGCATTCAATTATGTTCAGGTTCAGGGGAGGTGTGGGAGGTTTTAAAGCAAGTAAAA  
CCTCTACAAATGTGGTAAAATCGATAAGGATCAATTGGCTTCAGGTACCGTCACGATGTA  
GGTCACGGTCTCGAACGCCGGTGCAGGGTGCAGGGCGTGCCTGGCTCCCCGGCGCGT  
ACTCCACCTCACCCATCTGGTCATCATGATGAACGGGTGAGGTGGCGTAGTTGATCCG  
GCGAACGCGCGGCCAACGGGAAGCCCTGCCCTCGAAACCGCTGGCGCGGTGGTACGGT  
GAGCACGGACGTGCGACGGCGTCGGCGGGTGCAGGGATACGCCGGCAGCGTCAGGGTTCT  
CGACGGTCACGGCGGGCATGTCGACAGCGAATTGATCCGTCACCGATGCCCTGAGAGCC  
TTCAACCCAGTCAGCTCCTCCGGTGGCGGGGATGACTATCGTCGCCGACTTATGAC  
TGTCTTCTTATCATGCAACTCGTAGGACAGGTGCCGGCAGC

*Fig. 16*

**pRSV-C31int (SEQ ID NO: 9)**

CTGCATTAATGAATCGGCCAACCGCGGGGAGAGGCCGTTGCGTATTGGCGCTCTTCC  
GCTTCCTCGCTCACTGACTCGCTCGCTCGGCTGCGGCGAGCGGTATCAGCT  
CACTCAAAGCGGTAATACGGTTATCCACAGAATCAGGGATAACGCAGGAAAGAACATG  
TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGTTGCTGGCGTTTC  
CATAGGCTCCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA  
AACCCGACAGGACTATAAAGATACCAGGCCTTCCCCCTGGAAGCTCCCTCGCGCTCT  
CCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTCGGAAGCGTG  
GCGCTTCTCAATGCTACGCTGTAGGTATCTCAGTCGGTAGGTCGCTCGCTCCAAG  
CTGGCTGTGTGCACGAACCCCCCGTTCAGCCGACCGCTGCGCCTATCCGGTAACAT  
CGTCTTGAGTCCAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCCACTGGTAAC  
AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAAC  
TACGGCTACACTAGAAGGACAGTATTGGTATCTCGCCTGCTGAAGCCAGTTACCTTC  
GGAAAAAGAGTTGGTAGCTCTTGATCCGGAAACAAACCACCGCTGGTAGCGGTGGTTT  
TTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTGATC  
TTTCTACGGGCTGACGCTCAGTGGAACGAAAACACGTTAACGGATTGGTCA  
AGATTATCAAAAGGATCTCACCTAGATCCTTAAATTAAATGAAGTTAAATCA  
ATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA  
CCTATCTCAGCGATCTGCTATTGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA  
ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCCTGCAATGATACCGCGAGAC  
CCACGCTACCGGCTCCAGATTATCAGCAATAAACACAGCCAGCCGGAAAGGGCCAGCGC  
AGAAGTGGCCTGCAACTTATCCGCCTCCATCCAGTCTATTAAATTGTTGCCGGAAAGCT  
AGAGTAAGTAGTCGCCAGTTAATAGTTGCGAACGTTGCTTGCCATTGCTACAGGCATC  
GTGGTGTACGCTCGTGGTATGGCTTCACTCAGCTCCGGTCCAACGATCAAGG  
CGAGTTACATGATCCCCATGTTGCAAAAAGCGGTTAGCTCCTCGTCCTCCGATC  
GTTGTCAGAAGTAAGTGGCCGCAGTGTATCACTCATGGTTATGGCAGCAGTGCATAAT  
TCTCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTAGTACTCAACCAAG  
TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTGCCGGCGTCAATACGGAT  
AATACCGGCCACATAGCAGAACTTTAAAGTGCTCATCATTGAAAACGTTCTCGGGG  
CGAAAACCTCAAGGATCTTACCGCTGTTGAGATCCAGTTGATGTAACCCACTCGTGCA  
CCCAACTGATCTCAGCATCTTACCAACAGCTTCTGGGTGAGCAAAAACAGGA  
AGGCAAAATGCCGAAAAAGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC  
TTCCTTTCAATATTGAAAGCATTTACAGGGTTATTGTCATGAGCGGATACATA  
TTTGAATGTATTAGAAAAATAACAAATAGGGTCTCGCGCACATTCCCCGAAAAGTG  
CCACCTGACGTGACGGATCGGGAGATCTCCGATCCCCTATGGTCACTCTCAGTACAA  
TCTGCTCTGATGCCGATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTTGGAGGTG  
CTGAGTAGTGCAGCAGCAAAATTAAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCA  
TGAAGAATCTGCTTAGGGTTAGGCCTTGCCTGCGATGTAACGGGCCAGATATA  
CGCGTCTAGGGTCTAGGATCGATTAGGAATTCTCTAGCCGGTCTAGGGATCCCG  
GCGCGTATGGTGCAGTCTCAGTACAATCTGCTCTGATGCCGATAGTTAAGCCAGTATCT  
GCTCCCTGCTTGTGTTGGAGGTGCTGAGTAGTGCAGCAGCAAAATTAAAGCTACAAC  
AAGGCAAGGCTTGACCGACAATTGATGAAGAATCTGCTTAGGGTTAGGCCTTGCCT  
GCTTCGCGATGTACGGGCCAGATATACTCGCTATCTGAGGGACTAGGGTGTGTTAGGCG  
AAAAGCGGGCTCGGTGTACGCGTTAGGAGTCCCTCAGGATATACTGAGTCTCGCTT  
TTGCATAGGGAGGGGAAATGTAGTCTTATGCAATACTGAGTCTGCAACATGGTA  
ACGATGAGTTAGCAACATGCCTTACAAGGAGAGAAAAGCACCGTGCATGCCATTGGT  
GAAGTAAGGTGGTACGATCGTGCCTTATTAGGAAGGCAACAGACAGGTCTGACATGGATT  
GGACGAACCACTGAATTCCGCATTGCAAGAGATAATTGATTTAAGTGCCTAGCTCGATAC  
AATAAACGCCATTGACCATTACCACTTGGTGTGCACCTCCAAGCTGCATGCCATGCA  
GGTACCGGTCCGGAATTCCCGGGTCGACGAGCTCACTAGTCGTTAGGGCGACATGAC  
ACAAGGGGTTGTGACCGGGTGGACACGTACGCGGGTCTACGACCGTCAGTCGCGCGA

GCAGCGAGAATTGAGCGCAGCAAGCCCAGCGACACAGCGTAGCGCCAACGAAGACAAGGC  
GGCCGACCTTCAGCGCGAAGTCGAGCGCGACGGGGCCGGTCAGGTTCGTCGGCATT  
CAGCGAAGCGCCGGCACGTCGGCGTCGGACGGCGAGCGCCGGAGTCGAACGCAT  
CCTGAACGAATGCCGCCGGCGCTAACATGATCATTGTCTATGACGTGTCGCGCTT  
CTCGCGCCTGAAGGTATGGACGCGATTCCGATTGTCTCGGAATTGCTGCCCTGGCGT  
GACGATTGTTCACTCAGGAAGGCCTTCGGCAGGGAAACGTCATGGACCTGATTCA  
CCTGATTATGCGGCTCGACCGTGCACAAAGAATCTCGCTGAAGTCGGCGAAGATTCT  
CGACACGAAGAACCTTCAGCGCGAATTGGCGGGTACGTGGCGGGAAAGGCGCCTACGG  
CTTCGAGCTTGGTGGCGTGAGATCACGGCAACGGCCAATGGTCAATGTCGT  
CATCAACAAGCTGCGCACTCGACCACTCCCTTACCGGACCCCTCGAGTTGAGCCCGA  
CGTAATCCGGTGGTGGCGTGAGATCAAGACGACAAACACCTTCCCTCAAGCCGG  
CAGTCAAGCCGCCATTCACCCGGCAGCATCACGGGCTTGTAAAGCGCATGGACGCTGA  
CGCCGTGCCGACCCGGCGAGACGATTGGGAAGAAGACCGCTCAAGCGCCTGGACCC  
GGCAACCCTATGCGAATCCTCGGGACCCCGTATTGCGGGCTTCGCCGCTGAGGTGAT  
CTACAAGAAGAACCGGACGGCACGCCGACCACGAAGATTGAGGGTTACCGCATTGAGCG  
CGACCCGATCACGCTCCGGCCGGTCGAGCTTGATTGCGGACCGATCATCGAGCCGCTGA  
GTGGTATGAGCTTCAGGCGTGGTGGACGGCAGGGGGCGCGGCAAGGGGCTTCCGGGG  
GCAAGCCATTCTGTCCGCCATGGACAAGCTGTACTGCGAGTGTGGCGCCGTATGACTTC  
GAAGCGGGGAAGAACGATCAAGGACTCTTACCGCTGCCGTGCCGGAAAGGTGGTCGA  
CCCGTCCGCACCTGGGAGCACGAAGGCACGTGCAACGTACGATGGCGGCACTGACAA  
GTTCGTTGCCAACGCATCTCAACAAGATCAGGCACGCCGAAGGCGACGAAGAGACGTT  
GGCGCTCTGTGGGAAGCCGCCGACGCTCGGCAAGCTCACTGAGGCGCCTGAGAAGAG  
CGCGAACGGCGAACCTTGTGGCAGCGCCGACGCCCTGAACGCCCTGAAGAGAGCT  
GTACGAAGACCGCGCGGCAGGCGCGTACGACGGACCCGTTGGCAGGAAGCACTCCGGAA  
GCAACAGGCAGCGCTGACGCTCCGGAGCAAGGGGGAAAGAGCGGCTGCCGAACCTGA  
AGCCGCCGAAGCCCCGAAGCTCCCTTGACCAATGGTCCCCGAAGACGCCGACGCTGA  
CCCGACCGGCCCTAAGTCGTGGTGGGGCGCGCGTCAGTAGACGACAAGCGCGTGTGCGT  
CGGGCTTCGTTAGACAAGATCGTGTACGAAGTCGACTACGGCAGGGGGCAGGGAAC  
GCCCATCGAGAACCGCGCTTCGATCACGTGGCGAAGCCGCCGACCGACGACGAAGA  
CGACGCCAGGACGGCACGGAGACGTAGCGCGTAGCGAGACACCCGGATCCCTCGAGG  
GGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCCT  
TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCGTGCCTTCCTGACCCCTGGAAGGT  
GCCACTCCACTGTCCTTCTTAATAAAATGAGGAATTGCATCGCATTGTCTGAGTAGG  
TGTCAATTCTATTCTGGGGGTGGGTGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC  
AATAGCAGGCATGCTGGGATGCGGTGGCTCTATGGCTCTGAGGCGGAAAGAACCGAGG  
TGCCCCAGTCATAGCCGAATAGCCTCTCCACCAAGCGGCCGGAGAACCTGCGTGCATCC  
ACTGGGGCGCG

**Fig. 17**

**pCR-XL-TOPO-CMV-PUR-attB (SEQ ID NO: 10)**

AGCGCCCAATACGCAAACCGCCTCTCCCCGCGCTGGCCGATTCAATTATGCAGCTGGC  
ACGACAGGTTCCCGACTGGAAAGCGGGCAGTGAGCGAACGCAATTAAATGTGAGTTAGC  
TCACTCATTAGGCACCCAGGCTTACACTTATGCTCCGGCTCGTATGTTGTGGAA  
TTGTGAGCGGATAACAATTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTAT  
TTAGGTGACCGCGTTAGAATACTCAAGCTATGCATCAAGCTTGGTACCGAGCTCGGATCCA  
CTAGTAACGCCGCCAGTGTGCTGGAATTGCCCTGGCCGCAATAAAATATCTTATT  
TCATTACATCTGTGTTGGTTTGTGAAATCGATAGTACTAACATACGCTCTCCAT  
CAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGT  
GCCAGAACATTCTATCGATAGGTACCGAGCTTACGCGTGTAGCCCTCGAGCAGG  
ATCTATACATTGAATCAATATTGCAATTAGCCATATTAGTCATTGGTTATATAGCATAA  
ATCAATATTGGCTATTGCCATTGCATACGTTGTATCTATATCATAATATGTACATT  
ATTGGCTATGTCATATGACCGCCATGTTGACATTGATTATTGACTAGTTATTAA  
TAATCAATTACGGGGTCAATTAGTCATAGCCCATTATGGAGTTCCGCGTTACATAAC  
ACGGTAAATGGCCCGCTGGCTGACGCCAACGACCCCCGCCATTGACGTCAATGGGTGGAGT  
ACGTATGTTCCCAGTAGTAACGCCAATAGGGACTTCCATTGACGTCAATGGGTGGAGT  
TTACGGTAAACTGCCACTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCC  
ATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAGTACATGACCTTACGG  
GACTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG  
TTTGGCAGTACATCAATGGCGTGGTAGCGGTTGACTCACGGGATTCCAAGTC  
CACCCATTGACGTCAATGGAGTTGGCATTGGCACAAATCAACGGACTTCCAAA  
TGTGTAACAACCTCGCCCCATTGACGCAAATGGCGGTAGGCGTGTACGGTGGGAGGT  
TATATAAGCAGAGCTCGTTAGTGAACCGTCAGATGCCCTGGAGACGCCATCCACGCT  
TTTGACCTCCATAGAACGACACCGGGACCGATCCAGCCTCCCTCGAAGCTCGACTCTAG  
GGCTCGAGATCTGCGATCTAAGTAAGCTTGCATGCCCTGAGGTGGCCACGACCGGT  
GCCGCCACCATCCCTGACCCACGCCCTGACCCCTCACAGGAGACGACCTCCATGAC  
CGAGTACAAGCCCACGGTGCCTGCCACCGCACGACGTCCCCGGCGTACGCAC  
CCTGCCGCCCGTTCGCCACTACCCGCCACGCCACACCGTCGACCCGGACGCCA  
CATCGAGCGGGTACCGAGCTGCAAGAACCTTCTCACGCCGTGGCTCGACATCG  
CAAGGTGTGGTCGGGACGACGGGCCGGTGGCGAGATCGGCCGCGATGGCGAGTT  
CGAAGCGGGGGCGGTGTTGCCAGGAGATCGGCCGCGATGGCGAGTTGAGCGGTT  
GCTGGCCCGCAGCAAACAGATGGAAGGCCCTCTGGCGCCGACCGGGCCAAGGAG  
GTGGTTCTGGCCACCGTCGGCTCTGCCGACCCAGGGCAAGGGCTGGCAGCGC  
CGTCGTGCTCCCGGAGTGGAGGCCGAGCGCGCCGGGTGCCCTCCTGGAGAC  
CTCCGCCCGCAACCTCCCTTACGAGCGGCTCGGCTTACCGTCACGCCGACGT  
CGAGGTGCCGAAGGACCGCGCACCTGGTCATGACCCGAAGCCGGTGCCTGACGCC  
GCCACGCCCGCAGCGCCGACCGAACAGGAGCGCACGCCATGGCTCCACCGAAG  
CCGACCCGGCGGCCCGCCGACCCCGCACCGCCGGGAGGCCACCGACTCTAGAGTC  
GGGGCGGGCCGGCGCTTCGAGCAGACATGATAAGATAACATTGATGAGTT  
CAACTAGAATGCAGTGAACAAATGCTTATTGTGAAATTGTGATGCTATTGCTTAT  
TTGTAACCATTATAAGCTGCAATAAACAAAGTTAACACAACAAATTGCAATT  
TTCAGGTTCAAGGGGGAGGTGTGGGAGGTTTTAAAGCAAGTAAACCTCTACAAAT  
GTAAAATCGATAAGGATCAATTGGCTTCAGGTACCGTCACGATGTAGGT  
GAAGCCGGTGCAGGGTGCAGGGCGTGCCTGGCTCCCGGGCGCTACTCCAC  
ACCCATCTGGTCCATCATGATGAAACGGGTCAGGGTGGCGGTAGTTGAT  
GCCGCGCACCGGAAGCCCTGCCCTCGAAACCGCTGGCGGGTACGGTGAGC  
GGGACGTGCCACGGCGTCGGCGGATACCGGGGCGAGCGTCAGCGGGTCTCGAC  
GGTCACGGCGGGCATGTCGACAGCCGAATTGATCCGTCACCGATGCCCT  
CAACCCAGTCAGCTCCCGTGGCGGGCATGACTATCGTCGCCGACTTATGAC  
TGTCTCTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTCCGCT  
TCACTGACTCGCTCGCTCGGTGCTCGGCTGCCGAGCGGTATCAGCT  
CGGTAATACGGTTATCCACAGAATCAGGGATAACGCAGGAAAGAACAT  
TCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGG  
CCTATAGTGAAGTCGTATTACAATTCACTGGCCGTGTTACAAC  
GTGACTGGGAAACCGCTGGCGTACCGACATCCCCCTTCGCCAGCTGGCGTA  
ACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTCGCCAGCTGGCGTA

ATAGCGAAGAGGCCCGACCGATGCCCTCCCAACAGTTGCGCAGCCTATACGTACGGC  
AGTTAAGGTTACACCTATAAAAGAGAGAGGCCTTATCGTCTGTTGTGGATGTACAGA  
GTGATATTATTGACACGCCGGGGCACGGATGGTATCCCTGCCAGTGCACGTCTGC  
TGTCAAGATAAAAGTCTCCGTGAACTTACCCGGTGGCATATCGGGATGAAAGCTGGC  
GCATGATGACCACCGATATGCCAGTGTGCCGGTCTCCGTATCGGGGAAGAAGTGGCTG  
ATCTCAGCCACCGCGAAAATGACATCAAAACGCCATTAACCTGATGTTCTGGGAATAT  
AAATGTCAGGCATGAGATTATCAAAAAGGATCTCACCTAGATCCTTACGTAGAAAG  
CCAGTCAGCAAACGGTGTGACCCGGATGAATGTCACTACTGGCTATCTGGACAA  
GGGAAAACGCAAGCGAAAGAGAAAGCAGGTAGCTTGCAGTGGCTTACATGGCGATAGC  
TAGACTGGCGGTTTATGGACAGCAAGCGAACCGGAATTGCCAGCTGGGCGCCCTCTG  
GTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTCTCGCCGCCAAGGATCTGAT  
GGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTGCATGATTGAAAC  
AAGATGGATTGCACGCAGGTTCTCCGGCCCTGGGTGGAGAGGCTATTGGCTATGACT  
GGGCACAAACAGACAATCGGCTGCTTGATGCCGCGTGTCCGGCTGTCAGCGCAGGGG  
GCCCGGTTCTTTGTCAAGACCGACCTGTCGGTGCCTGAATGAACTGCAAGACGAGG  
CAGCGCGGCTATCGTGGCTGCCACGACGGCGTTGCAGCTGTGCTGACGTTG  
TCACTGAAGCGGGAAAGGGACTGGCTGCTATTGGCGAAGTGGCCGGGAGGATCTCCTGT  
CATCTCACCTTGCTCCTGCCAGAAAGTATCCATCATGGCTGATGCAATGCGGCGCTGC  
ATACGCTTGATCCGGTACCTGCCATTGACCAAGCGAAACATCGCATCGAGCGAG  
CACGTACTCGGATGGAAGCCGGTCTGATCAGGATGATCTGGACGAAGAGCATCAGG  
GGCTCGGCCAGCGAACTGTTGCCAGGCTCAAGGCAGCATGCCGACGGCGAGGATC  
TCGTCGTGACCCATGGCGATGCCTGCTGCCGAATATCATGGTGGAAATGGCCGCTTT  
CTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGG  
CTACCCGTGATATTGCTGAAGAGCTGGCCGGGAATGGCTGACCGCTTCTCGTGTGCTT  
ACGGTATGCCGCTCCGATTGCGCATCGCCTATCGCCTCTATCGCCTCTGACGAGTTCT  
TCTGAATTATTAACGCTTACAATTCTGATGCGGTATTTCTCCTTACGCATCTGTGCG  
GTATTTCACACCGCATACAGGTGGCACTTTGGGGAAATGTGCCGGAACCCCTATTG  
TTTATTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCCTGATAAAAT  
GCTTCAATAATAGCACGTGAGGAGGGCACCAGGCCAAGTGTGACCGAGTGGCTCCGGT  
GCTCACCGCGCGACGTGCCGGAGCGGGTGTGGCTGCCGGACGTGACCGCTGGTTCTC  
CGGGACTTCGTGGAGGACCGACTTCGCCGGTGTGGCTGCCGGACGTGACCGCTGGTCT  
CAGCGCGGTCCAGGACCAAGGTGGTGCAGAACACCCCTGGCTGGGTGTGGTGC  
CCTGGACGAGCTGTACGCCAGTGGTGGAGGTGTGACGAACTTCCGGACGCCTC  
CGGGCCGGCCATGACCGAGATCGCGAGCAGCCGTGGGGCGGAGTTGCCCTGCGCG  
CCCGGCCGGCAACTGCGTGACTTCGTGCCAGGAGCAGGACTGACACGTGCTAAACT  
TCATTTTAATTAAAAGGATCTAGGTGAAGAGATCCTTTGATAATCTCATGACCAAAAT  
CCCTTAACGTGAGTTCTGTTCCACTGAGCGTCAGACCCGTAGAAAAGATCAAAGGATC  
TTCTTGAGATCCTTTCTGCGCGTAATCTGCTGCTGCAAACAAAAAACCCACCGCT  
ACCAGCGGTGGTTGGATCAAGAGCTACCAACTCTTCCGAAGGTAACCTGG  
CTTCAGCAGAGCGCAGATAACAAATACTGTCCTCTAGTGTAGCCGTAGTTAGGCCACCA  
CTTCAAGAACTCTGTAGCACCGCCTACATACTCGCTCTGCTAACCTGTTACCGTGGC  
TGCTGCCAGTGGCGATAAGCGTGTCTACCGGGTTGGACTCAAGACGATAGTTACCGGA  
TAAGGCGAGCGGTGGCTGAACGGGGGGTGTGACACAGCCCAGCTGGAGCGAAC  
GACCTACACCGAAGTGGATACCTACAGCGTGAGCTATGAGAAAGGCCACGCTCCGA  
AGGGAGAAAGGCCAGGTATCCGTAAGCGGCAGGGTGGAAACAGGAGAGCGCACGAG  
GGAGCTCCAGGGGAAACGCCCTGGTATCTTATAGTCCTGTCGGGTTGCCACCTCTG  
ACTTGAGCGTCGATTGGATGCTGTCAGGGGGCGGAGCCTATGGAAAAACGCCAG  
CAACGCCCTTTACGGTCTGGCTTTGCTGGCCTTGTGTCACATGTTCTTCC  
TGCCTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTGAGTGAGCTGATACCGC  
TCGCCGAGCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAG

37/39

**attP (SEQ ID NO: 11)**

GAAGTAGTACTGACGGACACACCGAAGCCCCGGCGGCAACCCTCAGCGGATGCCCGGGGCTT  
CACGTTTCCAGGTAGAAGCGGTTTGGAGTAGTGCCCCACTGGGTAACCTTGAG  
TTCTCTCAGTTGGGGCGTAGGGTCGCCGACATGACACAAGGGTTGTGACCGGGTGGACA  
CGTACGCGGGTGCTTACGACCGTCAGTCGCGAGCGCGACTAGTACA

*Fig. 19*

38/39

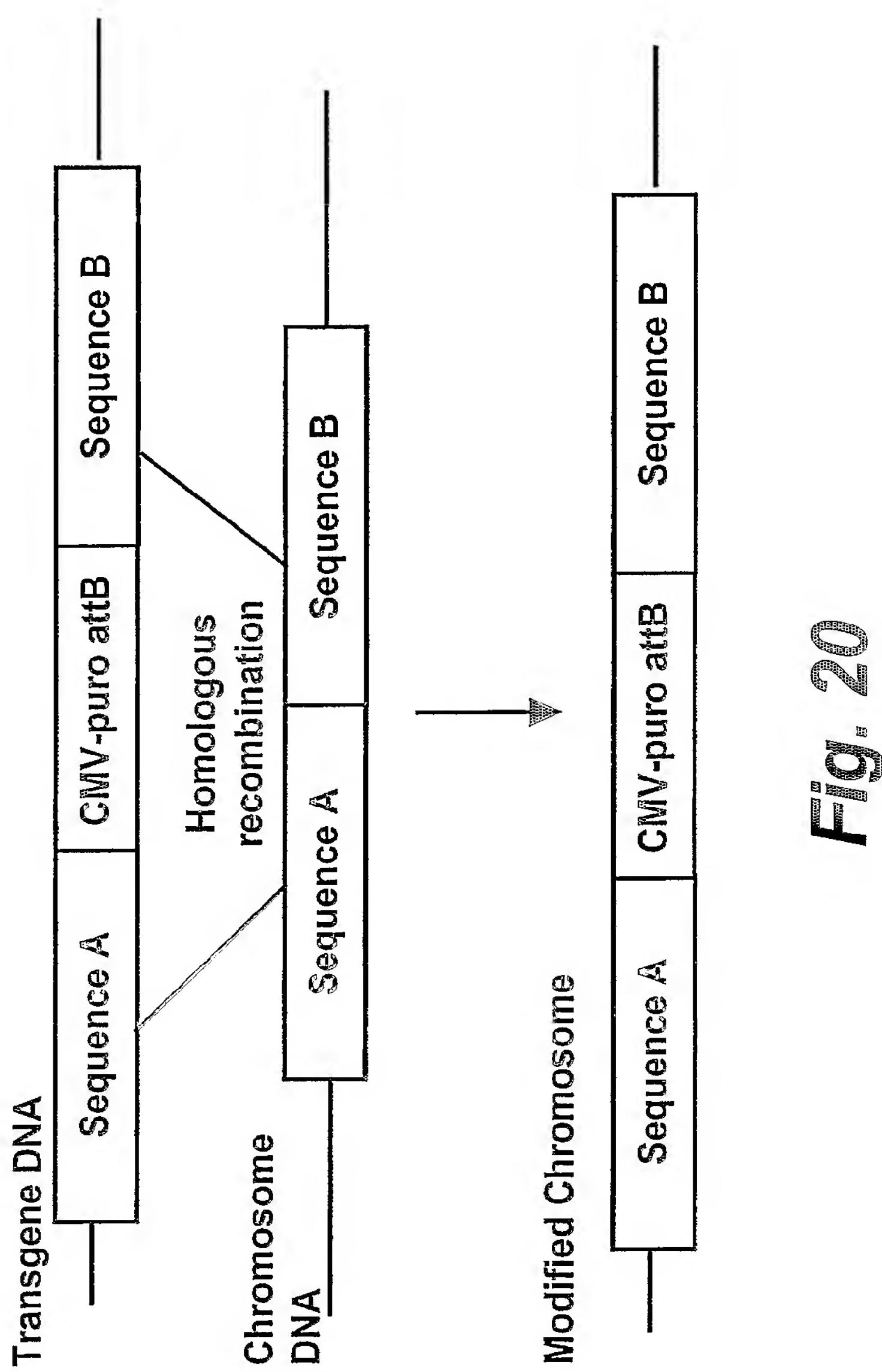


Fig. 20

39/39

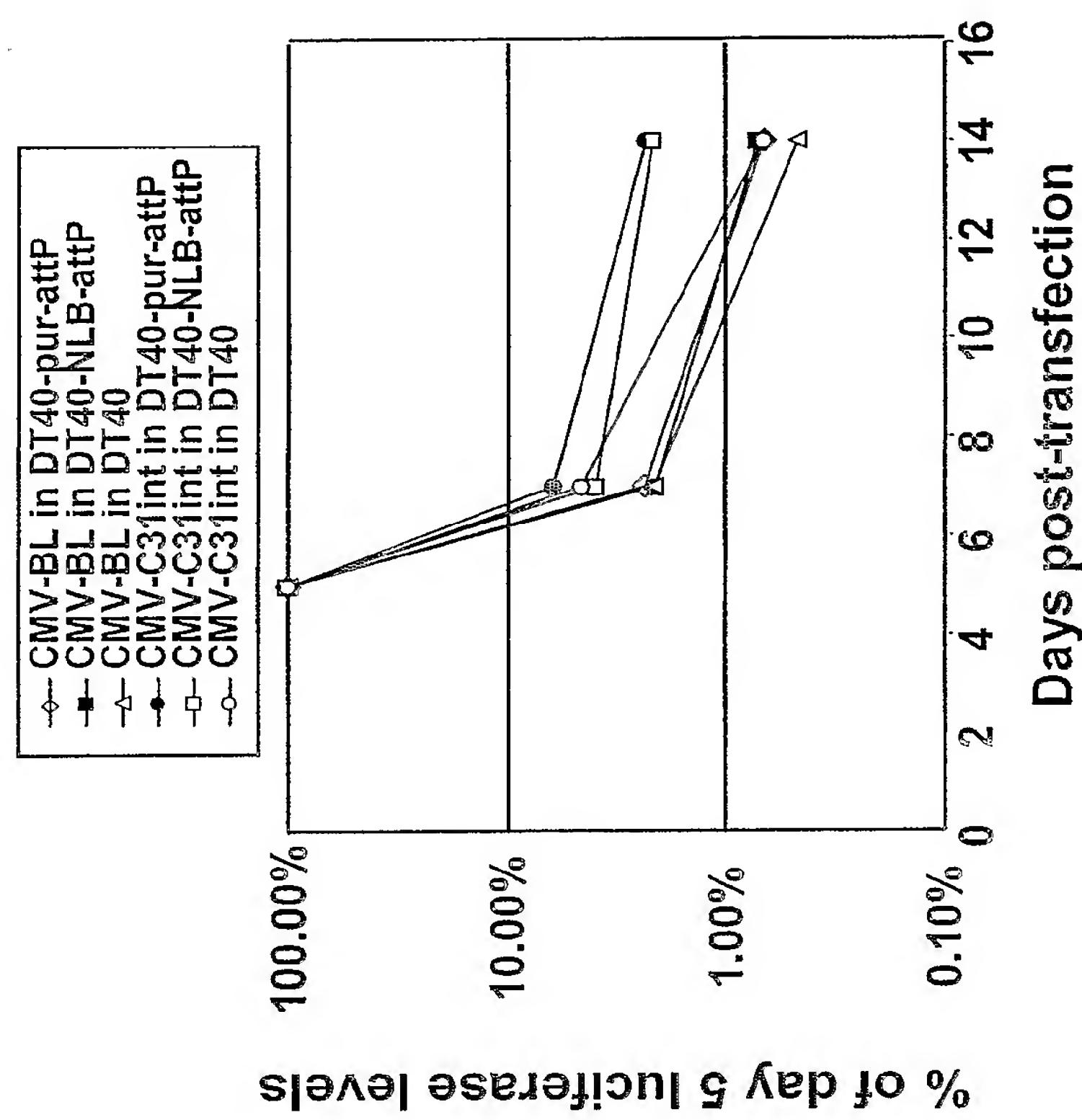


Fig. 21

Integrase.ST25  
SEQUENCE LISTING

<110> AviGenics, Inc

<120> Avian Integrase-mediated Transformation

<130> A181 1080.1

<160> 12

<170> PatentIn version 3.2

<210> 1

<211> 6230

<212> DNA

<213> Plasmid pCMV-31int

<400> 1

cattcgccat tcaggctgcg caactgttgg gaagggcgat cggcgcggc ctcttcgcta 60  
ttacgccagc caatacgcaa accgcctctc cccgcgcgtt ggccgattca ttaatgcagg 120  
atcgatccag acatgataag atacattgtat gagtttggac aaaccacaac tagaatgcag 180  
tgaaaaaaaaat gctttatttg tgaaatttgt gatgctatttgc ctttatttgt aaccattata 240  
agctgcaata aacaagttaa caacaacaat tgcattcatt ttatgtttca ggttcagggg 300  
gaggtgtggg aggttttta aagcaagtaa aacctctaca aatgtggat ggctgattat 360  
gatcatgaac agactgtgag gactgagggg cctgaaatga gccttgggac tgtgaatcta 420  
aaatacaca acaatttagaa tcactagctc ctgtgtataa tattttcata aatcataactc 480  
agtaagcaaa actctcaagc agcaagcata tgcagctatgt ttaacacatt atacacttaa 540  
aaattttata ttacacctt tagctttaaat ctctgttaggt agtttgtcca attatgtcac 600  
accacagaag taaggttcct tcacaaagat cccaaagctat cttataatac gactcactat 660  
aggagagag ctatgacgac gcatgcacgc gtaagcttgg gcccctcgag ggatccgggt 720  
gtctcgctac gcccgtacgt cttccgtgcc gtcctggcg tcgtcttcgt cgtcgctgg 780  
cgccggcttc gcccacgtga tcgaagcgcg cttctcgatg ggcgttccct gccccctgcc 840  
cgtagtcgac ttctgtacaa cgatcttgc tacgaagagc ccgacgaaca cgcgttgc 900  
gtctactgac gcgcccccc accacgactt agggccggc gggtcagcgt cggcgttttc 960  
ggggaaaccat tggtaaaggga aagcttcgg ggcttcggcg gcttcaagtt cggcaagccg 1020  
ctcttcggcc ctttgctgcc ggagcgtcag cgctgcctgt tgcttccgga agtgccttc 1080  
gccaacgggt ccgtcgatcg cgcctgccgc gcggcttcg tacagctttt caagggcg 1140  
cagggcgctcg gcgcgctccg caacaagggtt cgcccttcg ccgctttt cagggcgctc 1200  
agttagcttgc ccgaagcgac gggcgcttc ccacagaagc gccaacgtct cttcgctgcc 1260  
ttcggcgatgc ctgatcttgt tgaagatgcg ttccgcaacg aacttgcga gtgccgcatt 1320  
gctgacgttg cacgtgcctt cgtgcgtccc aggtgcggac gggtcgacca cttccggcg 1380  
acggcagcgg taagagtccct tgatcgattt ttccccgcgc ttcgaagtca tgacggcgcc 1440  
acactcgac tacagcttgtt ccatggcgat cagaatggct tgccccggg aaagccccctt 1500

Integrase.ST25

gccgcgcccc	ctgcccgtcca	accacgcctg	aagctcatac	cactcagcgg	gctcgatgat	1560
cggccgcaaa	tcaagctcga	ccggccggag	cgtgatcggg	tcgcgctgaa	tgcggtaacc	1620
ctcaatcttc	gtggtcggcg	tgccgtccgg	cttcttcttg	tagatcacct	cagcggcgaa	1680
gcccgcataa	cgcgggtccc	gaaggattcg	cataacggtt	gccgggtccc	aggcgcttga	1740
agcggtcttc	ttcccaatcg	tctcgccccg	ggtcggcacg	gcgtcagcgt	ccatgcgctt	1800
acaaagcccc	gtgatgctgc	ccgggtgaat	ggcggcttga	ctgcccggct	tgaagggaaag	1860
gtgtttgtgc	gtcttgatct	cacgcccacca	ccaccggatt	acgtcgggct	cgaactcgaa	1920
gggtccggta	aggggagtgg	tcgagtgcgc	aagcttggtg	atgacgacat	tgaccattcg	1980
gccgttgcgc	gtgatctcct	tcgtctccga	aacaagctcg	aagccgtaag	gcgccttccc	2040
gccgacgtac	ccgccccatt	cgcgctgaag	gttcttcgtg	tcgagaatct	tcgcccactt	2100
cagcgaagat	tctttgtcg	acgcgtcgag	ccgcataatc	aggtgaatca	ggtccatgac	2160
gtttccctgc	cggaaagacgc	cttcctgagt	ggaaacaatc	gtcacgccc	ggcgagcaa	2220
ttccgagaca	atcggaatcg	cgtccatgac	cttcaggcgc	gagaagcgcg	acacgtcata	2280
gacaatgatc	atgttgagcc	gcccggcgcg	gcattcggt	aggatgcgtt	cgaactccgg	2340
gcgctccgcc	gtcccgaacg	ccgacgtgcc	cggcgcttcg	ctgaaatgcc	cgacgaacct	2400
gaaccggccc	ccgtcgcgct	cgacttcgcg	ctgaaggtcg	gccgccttgc	cttcgttggc	2460
gctacgctgt	gtcgctggc	ttgctgcgt	cgaattctcg	cgctcgcg	actgacggtc	2520
gtaagcaccc	gcgtacgtgt	ccaccccggt	cacaacccct	tgtgtcatgt	cggcgaccct	2580
acgactagtg	agctcgtcga	cccggaaatt	ccggaccggt	acctgcaggc	gtaccttcta	2640
tagtgtcacc	taaatagctt	tttgcaaaag	cctaggctag	agtccggagg	ctggatcggt	2700
cccggtgtct	tctatggagg	tcaaaacagc	gtggatggcg	tctccaggcg	atctgacggt	2760
tcactaaacg	agctctgctt	atatacacct	cccaccgtac	acgcctaccg	cccatttgcg	2820
tcaatggggc	ggagtttta	cgacattttg	gaaagtcccg	ttgattttgg	tgc当地aca	2880
aactcccatt	gacgtcaatg	gggtggagac	ttggaaatcc	ccgtgagtca	aaccgctatc	2940
cacgcccatt	gatgtactgc	caaaaccgca	tcaccatgg	aatagcgatg	actaatacg	3000
agatgtactg	ccaagttagga	aagtccata	aggtcatgta	ctggcataa	tgccaggcgg	3060
gccatttacc	gtcattgacg	tcaatagggg	gcgtacttgg	catatgatac	acttgatgta	3120
ctgccaagtg	ggcagttac	cgtaaatact	ccacccattg	acgtcaatgg	aaagtcccta	3180
ttggcgttac	tatgggaaca	tacgtcatta	ttgacgtcaa	tggcgcccc	tcgttggcgc	3240
gtcagccagg	cgggcccattt	accgtaagtt	atgtaacgac	ctgcacgatg	ctgtttcctg	3300
tgtgaaattg	ttatccgctc	acaattccac	acattatacg	agccggaagc	tataaagtgt	3360
aaagcctggg	gtgcctaattg	agtgaaaggg	cctcgatatac	gcctattttt	ataggttaat	3420
gtcatgataa	taatggtttc	ttagacgtca	ggtggcactt	ttcggggaaa	tgtgcgcgaa	3480
accctattt	gtttattttt	ctaaatacat	tcaaataatgt	atccgctcat	gagacaataa	3540

Integrase\_ST25

ccctgataaa	tgcttcaata	atattgaaaa	acgcgcgaat	tgcaagctct	gcattaatga	3600
atcgcccaac	gcgcggggag	aggcggttg	cgtattggc	gctttccgc	ttcctcgctc	3660
actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	3720
gtaatacggt	tatccacaga	atcagggat	aacgcaggaa	agaacatgtg	agcaaaaggc	3780
cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgttttcca	taggctccgc	3840
ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccgacagga	3900
ctataaagat	accaggcggtt	tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	3960
ctgccgctta	ccggataacct	gtccgcctt	ctcccttcgg	gaagcgtggc	gctttctcaa	4020
tgctcacgct	gtaggtatct	cagttcggtg	taggtcgttc	gctccaagct	gggctgtgtg	4080
cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	4140
aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	4200
gcgaggtatg	taggcggtgc	tacagagttc	ttgaagtgg	ggcctaacta	cggctacact	4260
agaaggacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	4320
ggtagcttt	gatccggcaa	acaaaccacc	gctggtagcg	gtggttttt	tgttgcaag	4380
cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	4440
tctgacgctc	agtggAACGA	aaactcacgt	taaggattt	tggtcatgcc	ataacttcgt	4500
atagcataca	ttatacgaag	ttatggcatg	agattatcaa	aaaggatctt	cacctagatc	4560
cttttaaatt	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aacttggct	4620
gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	atttcgttca	4680
tccatagttg	cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	cttaccatct	4740
ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	tttatcagca	4800
ataaaccagc	cagccggaag	ggccgagcgc	agaagtggc	ctgcaacttt	atccgcctcc	4860
atccagtcta	ttaattgtt	ccggaaagct	agagtaagta	gttcgccagt	taatagttt	4920
cgcaacgttg	ttgccattgc	tacaggcatc	gtgggtcac	gctcgtcg	tggatggct	4980
tcattcagct	ccggttccca	acgatcaagg	cgagttacat	gatccccat	gttgtgcaaa	5040
aaagcggta	gctccttcgg	tcctccgatc	gttgtcagaa	gtaagttggc	cgcagtgtta	5100
tcactcatgg	ttatggcagc	actgcataat	tctttactg	tcatgccatc	cgtaagatgc	5160
ttttctgtga	ctgggtgagta	ctcaaccaag	tcaattctgag	aatagtgtat	gcggcgaccg	5220
agtgtcttt	gcccggcgtc	aatacggat	aataccgcgc	cacatagcag	aactttaaaa	5280
gtgctcatca	ttggaaaacg	ttttcgggg	cgaaaactct	caaggatctt	accgctgttg	5340
agatccagtt	cgtatgtacc	cactcggtca	cccaactgat	cttcagcatc	ttttactttc	5400
accagcgttt	ctgggtgagc	aaaaacagga	aggcaaaatg	ccgcaaaaaaa	ggaaataagg	5460
gacacacgga	aatgttgaat	actcatactc	ttcctttttc	aatattattg	aagcatttat	5520
cagggttatt	gtctcatgcc	aggggtggc	acacatattt	gataccagcg	atccctacac	5580

Integrase\_ST25

agcacataat	tcaatgcgac	ttccctctat	cgcacatctt	agacctttat	tctccctcca	5640
gcacacatcg	aagctgccga	gcaagccgtt	ctcaccagtc	caagacctgg	catgagcgga	5700
tacatatttgc	aatgtattta	gaaaaataaa	caaatagggg	ttccgcgcac	atttccccga	5760
aaagtgccac	ctgaaattgt	aaacgttaat	atttgttaa	aattcgcgtt	aaattttgt	5820
taaatcagct	catttttaa	ccaataggcc	gaaatcgga	aaatccctta	taaatcaaaa	5880
gaatagaccg	agatagggtt	gagtgttgg	ccagttgg	acaagagtcc	actattaaag	5940
aacgtggact	ccaacgtcaa	agggcgaaaa	accgtctatc	agggcgatgg	cccactacgt	6000
gaaccatcac	cctaatacg	tttttgggg	tcgaggtgcc	gtaaagcact	aaatcggAAC	6060
cctaaaggga	cccccgatt	tagagcttga	cggggaaagc	cggcgaacgt	ggcgagaaag	6120
gaagggaaaga	aagcgaaagg	agcggcgct	agggcgctgg	caagtgttagc	ggtcacgctg	6180
cgcgtAACCA	ccacaccgc	cgcgttaat	gcccgcgtac	agggcgcg	tc	6230

<210> 2  
<211> 5982  
<212> DNA  
<213> Plasmid pCMV-Luc-attB

<400> 2	ctctatcgat	aggtaccgag	ctcttacgcg	tgctagccct	cgagcaggat	ctatacatttgc	60
	aatcaatatt	ggcaatttgc	catattagtc	attggttata	tagcataaat	caatattggc	120
	tattggccat	tgcatacg	tttgcgtt	gtatctat	cataatatgt	acatttat	180
	ccaatatgac	cgcgcgtt	acattgatta	ttgactagtt	attaatagta	atcaattacg	240
	gggtcattag	ttcatagccc	atatatggag	ttccgcgtt	cataacttac	ggtttatggc	300
	ccgcctggct	gaccgccc	cgaccccccgc	ccattgacgt	caataatgac	gtatgttccc	360
	atagtaacgc	caataggac	tttccattga	cgtcaatggg	tggagtattt	acggtaaact	420
	gccca	tttgcgtt	catatca	agtgtatcat	atgccaagtc	cgccccc	480
	gacggtaat	ggccgcctg	gcattatg	cgtacatga	ccttacgg	ctttcctact	540
	tggcagtaca	tctacgtatt	agtcatcg	attaccatgg	tgtgcgtt	ttggcagtac	600
	atcaatggc	gtggatagcg	gtttgactca	cggggatttc	caagtctcca	ccccattgac	660
	gtcaatggg	gtttgtttg	gcaccaaaat	caacggact	ttccaaaatg	tcgtacaac	720
	tccgc	ccat	tgacgcaat	ggcggtagg	cgtgtacgg	gggaggtcta	780
	gctcg	tttag	tgaaccgtca	gatgcctgg	agacgc	cacgctgtt	840
	agaagacacc	gggaccgatc	cagcctccc	tcgaagctcg	actctaggg	ctcgagatct	900
	gcgatctaag	taagcttggc	attccgg	tgttggtaaa	gccaccatgg	aagacgcca	960
	aaacataaa	aaaggcccgg	cgccattcta	tccgctggaa	gatgaaaccg	ctggagagca	1020
	actgcataa	gctatgaaga	gatacgcc	gttcc	acaattgctt	ttacagatgc	1080
	acatatcgag	gtggacatca	cttacgt	gtacttcgaa	atgtccgtt	ggttggcaga	1140
	agctatgaaa	cgatatggc	tgaatac	tcacagaatc	gtcgtatgca	gtgaaaactc	1200

## Integrase.ST25

tcttcatttc	tttatgcgg	tgttggcgc	gttatttac	ggagttgcag	ttgcgcgc	1260			
gaacgacatt	tataatgaac	gtgaattgct	caacagtatg	ggcatttcgc	agcctaccgt	1320			
ggtgttcgtt	tccaaaaagg	ggttgcaaaa	aattttgaac	gtgcaaaaaa	agctcccaat	1380			
catccaaaaa	attattatca	tggattctaa	aacggattac	cagggatttc	agtcgatgta	1440			
cacgttcgtc	acatctcatc	tacctccgg	tttaatgaa	tacgatttg	tgccagagtc	1500			
cttcgatagg	gacaagacaa	ttgcactgat	catgaactcc	tctggatcta	ctggctcgcc	1560			
taaaggtgtc	gctctgcctc	atagaactgc	ctgcgtgaga	ttctcgcatg	ccagagatcc	1620			
tatTTTggc	aatcaaatca	ttccggatac	tgcgattta	agtgttgttc	cattccatca	1680			
cgttttgg	atgtttacta	cactcggata	tttgatatgt	ggatttcgag	tcgtcttaat	1740			
gtatagattt	gaagaagac	tgTTTctgag	gagccttcag	gattacaaga	ttcaaagtgc	1800			
gctgctgg	ccaaccctat	tctcTTTctt	cgcAAaaagc	actctgattt	acaaatacga	1860			
tttatcta	ttacacgaaa	ttgcttctgg	tggcgctccc	ctctctaagg	aagtccggg	1920			
agcggttgc	aagaggttcc	atctgccagg	tatcaggcaa	ggatatggc	tcactgagac	1980			
tacatcag	attctgatta	cacccgaggg	ggatgataaa	ccggcgcgg	tcggtaaagt	2040			
tgttccattt	tttgaagcga	aggTTTgtgg	tctggatacc	gggAAAACgc	tggcgtaaa	2100			
tcaaagaggc	gaactgtgt	tgagagg	tatgattatg	tccggttatg	taaacaatcc	2160			
ggaagcgacc	aacgccttga	ttgacaagga	tggatggcta	cattctggag	acatagctta	2220			
ctgggacgaa	gacgaacact	tcttcatcgt	tgaccgcctg	aagtctctga	ttaagtacaa	2280			
aggctatcag	gtggctccc	ctgaatttga	atccatcttgc	ctccaacacc	ccaacatctt	2340			
cgacgcaggt	gtcgcagg	ttcccga	tgacgcccgt	gaactccc	ccgcccgtgt	2400			
tgttttggag	cacggaaaga	cgatgacg	aaaagagatc	gtggattacg	tcgcccagtca	2460			
agtaacaacc	gcgaaaaagt	tgcgcgagg	agttgtgtt	gtggacgaa	taccgaaagg	2520			
tcttaccgga	aaactcgac	caagaaaaat	cagagagatc	ctcataaaagg	ccaagaagg	2580			
cggaaagatc	gccgtgtaat	tctagagt	gggcggccgg	ccgcttcgag	cagacatgt	2640			
aagatacatt	gatgagtttgc	gacaaaccac	aactagaatg	cagtaaaaaa	aatgctttat	2700			
ttgtgaaatt	tgtgatgct	ttgcttatt	tgttaaccatt	ataagctgca	ataaacaagt	2760			
taacaacaac	aattgcattt	atTTTatgtt	tcaggttgc	ggggaggtgt	gggaggttt	2820			
ttaaagcaag	taaaacctct	acaaatgtgg	taaaatcgat	aaggatcaat	tcggcttcag	2880			
gtaccgtcga	cgtgttagt	cacggtctcg	aagccgcgt	gcgggtgcca	ggcgtgccc	2940			
ttgggctccc	cggcgcgt	ctccacctca	ccatctgtt	ccatcatgat	gaacgggtcg	3000			
aggTggcgt	agttgatccc	ggcgaacgcg	cggcgcaccg	ggaagccctc	gccctcgaaa	3060			
ccgctggcg	cgtgg	gtgagc	ggacgtgcga	cggcgtcggc	gggtgcggat	3120			
acgcggggca	gcgtcagcgg	gttctcgac	gtcacggcgg	gcatgtcgac	agccgaattt	3180			
atccgtc	cgatgcc	ttt	gagac	ttc	aacc	gtca	ttccg	gtggcgcgg	3240

## Integrase.ST25

ggcatgacta	tcgtcgccgc	acttatgact	gtcttcttta	tcatgcaact	cgttaggacag	3300
gtgccggcag	cgctcttccg	cttcctcgct	cactgactcg	ctgcgctcgg	tcgttcggct	3360
gcggcgagcg	gtatcagctc	actcaaaggc	ggtataacgg	ttatccacag	aatcagggga	3420
taacgcagga	aagaacatgt	gagcaaaagg	ccagcaaaag	gccaggaacc	gtaaaaaggc	3480
cgcgttgctg	gcgttttcc	ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	3540
ctcaagtcag	aggtggcgaa	acccgacagg	actataaaga	taccaggcgt	ttccccctgg	3600
aagctccctc	gtgcgctctc	ctgttccgac	cctgccgctt	accggatacc	tgtccgcctt	3660
tctcccttcg	ggaagcgtgg	cgctttctca	atgctcacgc	tgttaggtatc	tcagttcggt	3720
gtaggtcgtt	cgctccaagc	tgggctgtgt	gcacgaaccc	cccgttcagc	ccgaccgctg	3780
cgccttatcc	ggttaactatc	gtcttgagtc	caacccggta	agacacgact	tatcgccact	3840
ggcagcagcc	actggtaaca	ggattagcag	agcgaggtat	gtaggcggtg	ctacagagtt	3900
cttgaagtgg	tggcctaact	acggctacac	tagaaggaca	gtatttggta	tctgcgctct	3960
gctgaagcca	gttaccttcg	aaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	4020
cgctggtagc	ggtggtttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	4080
tcaagaagat	ccttgatct	tttctacggg	gtctgacgct	cagtggAACG	aaaactcact	4140
ttaagggatt	ttggcatga	gattatcaa	aaggatctc	acctagatcc	ttttaaatta	4200
aaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	acttggtctg	acagttacca	4260
atgcttaatc	agtgaggcac	ctatctcagc	gatctgtcta	tttcgttcat	ccatagttgc	4320
ctgactcccc	gtcgtgtaga	taactacgt	acgggagggc	ttaccatctg	gccccagtgc	4380
tgcaatgata	ccgcgagacc	cacgctcacc	ggctccagat	ttatcagcaa	taaaccagcc	4440
agccggaagg	gccgagcgca	gaagtggtcc	tgcaacttta	tccgcctcca	tccagtctat	4500
taattgttgc	cggaaagcta	gagtaagttag	ttcgccagtt	aatagttgc	gcaacgttgc	4560
tgccattgct	acaggcatcg	tgggtcactg	ctcgtcggtt	ggtatggctt	cattcagctc	4620
cggttcccaa	cgtcaaggc	gagttacatg	atccccatg	ttgtcaaaa	aagcggttag	4680
ctccttcggt	cctccgatcg	ttgtcagaag	taagttggcc	gcagtgttat	cactcatggt	4740
tatggcagca	ctgcataatt	ctcttactgt	catgccatcc	gtaagatgct	tttctgtgac	4800
tggtagtac	tcaaccaagt	cattctgaga	atagtgtatg	cggcgaccga	gttgccttt	4860
ccggcgtca	atacggata	ataccgcgcc	acatagcaga	actttaaaag	tgctcatcat	4920
tggaaaacgt	tttcggggc	gaaaactctc	aaggatcttta	ccgctgttga	gatccagttc	4980
gatgtAACCC	actcgtgcac	ccaaactgatc	ttcagcatct	tttactttca	ccagcggttc	5040
tgggtgagca	aaaacaggaa	ggcaaaatgc	cgaaaaaaag	ggaataaggg	cgacacggaa	5100
atgttgaata	ctcataactct	tccttttca	atattattga	agcatttatac	agggttattt	5160
tctcatgagc	ggatacatat	ttgaatgtat	tttagaaaaat	aaacaaatag	gggtccgcg	5220
cacatttccc	cgtaaagtgc	cacctgacgc	gccctgttagc	ggcgcattaa	gcgccggcggg	5280

## Integrase.ST25

tgtgggtggtt acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt	5340
cgccttcttc ctttcctttc tcgccacgtt cgccggctt ccccgtaag ctctaaatcg	5400
ggggctccct ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga	5460
ttagggtat ggttcacgta gtgggccatc gccctgatag acggttttc gccctttgac	5520
gttggagtcc acgttctta atagtggact cttgttccaa actggaacaa cactcaaccc	5580
tatctcggtc tattctttg atttataagg gatttgccc atttcggcct attggtaaa	5640
aatgagctg atttaacaaa aatttaacgc gaatttaac aaaatattaa cgttacaat	5700
ttcccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc gggcctttc	5760
gctattacgc cagcccaagc taccatgata agtaagtaat attaaggtac gggaggtact	5820
tggagcggcc gcaataaaaat atctttattt tcattacatc tgtgtgttgg tttttgtgt	5880
gaatcgatag tactaacata cgctctccat caaaacaaaaa cgaaacaaaaa caaactagca	5940
aaataggctg tccccagtc aagtgcaggt gccagaacat tt	5982

<210> 3  
<211> 5924

<212> DNA

<213> Plasmid pCMV-Luc-attP

<400> 3 ctctatcgat agtaccgag ctcttacgcg tgcttagccct cgagcaggat ctatacattt	60
aatcaatatt ggcaatttgc catattagtc attggttata tagcataaat caatattggc	120
tattggccat tgcatacggt gtatctatat cataatatgt acatttatat tggctcatgt	180
ccaatatgac cgccatgtt acattgatta ttgacttagtt attaatagta atcaattacg	240
gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac ggtttatggc	300
ccgcctggct gaccgccaa cgaccccccgc ccattgacgt caataatgac gtatgttccc	360
atagtaacgc caataggac tttccattga cgtcaatggg tggagtattt acggtaaact	420
gcccaacttgg cagtacatca agtgtatcat atgccaagtc cgccccctat tgacgtcaat	480
gacggtaaat ggccgcctg gcattatgcc cagtacatga ctttacggga ctttcctact	540
tggcagtaca tctacgtatt agtcatcgct attaccatgg tggatgcgtt ttggcagtac	600
atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca ccccattgac	660
gtcaatggga gtttgggttg gcacaaaaat caacggact ttccaaaaatg tcgtaacaac	720
tccgccccat tgacgcaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga	780
gctcggttag tgaaccgtca gatgcctgg agacgccatc cacgctgttt tgacccat	840
agaagacacc gggaccgatc cagcctcccc tcgaagctcg actctagggg ctcgagatct	900
gcgtatctaag taagcttggc attccggtagc tggatggtaaa gccaccatgg aagacgcca	960
aaacataaaag aaaggcccg cgccattcta tccgctggaa gatggaaaccg ctggagagca	1020
actgcataag gctatgaaga gatacgcct gtttcctggaa acaattgctt ttacagatgc	1080

Integrase.ST25

acatatcgag	gtggacatca	cttacgctga	gtacttcgaa	atgtccgttc	ggttggcaga	1140
agctatgaaa	cgatatggc	tgaatacataa	tcacagaatc	gtcgtatgca	gtgaaaactc	1200
tcttcaattc	tttatgccgg	tgttggcgc	gttatttatac	ggagttgcag	ttgcgcggc	1260
gaacgacatt	tataatgaac	gtgaattgct	caacagtatg	ggcatttcgc	agcctaccgt	1320
ggtgttcgtt	tccaaaaagg	ggttgcaaaa	aattttgaac	gtgaaaaaaa	agctcccaat	1380
catccaaaaa	attattatca	tggattctaa	aacggattac	cagggatttc	agtcgatgta	1440
cacgttcgtc	acatctcatc	tacctccgg	tttaatgaa	tacgattttg	tgccagagtc	1500
cttcgatagg	gacaagacaa	ttgcactgat	catgaactcc	tctggatcta	ctggctgccc	1560
taaaggtgtc	gctctgcctc	atagaactgc	ctgcgtgaga	ttctcgcatg	ccagagatcc	1620
tatTTTggc	aatcaaatac	ttccggatac	tgcgatttt	agtgttgc	cattccatca	1680
cggTTTgga	atgtttacta	cactcgata	tttgatatgt	ggatttcgag	tcgtcttaat	1740
gtatagattt	gaagaagagc	tgtttctgag	gagccttcag	gattacaaga	ttcaaagtgc	1800
gctgctggtg	ccaaccctat	tctccttctt	cgc当地aaagc	actctgattt	acaaatacga	1860
tttatctaatt	ttacacgaaa	ttgcttctgg	tggcgctccc	ctctctaagg	aagtcgggga	1920
agcggttgc	aagaggttcc	atctgccagg	tatcaggcaa	ggatatggc	tcactgagac	1980
tacatcagct	attctgatta	cacccgaggg	ggatgataaa	ccggcgcgg	tcggtaaagt	2040
tgttccattt	tttgaagcga	aggttgcgaa	tctggatacc	gggaaaacgc	tggcgtaaa	2100
tcaaagaggc	gaactgtgt	tgagaggtcc	tatgattatg	tccggttatg	taaacaatcc	2160
ggaagcgacc	aacgccttga	ttgacaagga	tggatggcta	cattctggag	acatagctta	2220
ctgggacgaa	gacgaacact	tcttcatcgt	tgaccgcctg	aagtctctga	ttaagtacaa	2280
aggctatcag	gtggctccc	ctgaatttga	atccatcttgc	ctccaacacc	ccaacatctt	2340
cgacgcaggt	gtcgcaggc	ttcccgacga	tgacgcccgt	gaacttccc	ccgcccgttgc	2400
tgtttggag	cacggaaaga	cgatgacgga	aaaagagatc	gtggattacg	tcgccagtca	2460
agtaacaacc	gc当地aaaaagt	tgcgcggagg	agttgtgtt	gtggacgaag	taccgaaagg	2520
tcttaccgga	aaactcgacg	caagaaaaat	cagagagatc	ctcataaagg	ccaagaaggg	2580
cggaaagatc	gccgtgtaat	tctagagtcg	gggcggccgg	ccgcttcgag	cagacatgat	2640
aagatacatt	gatgagtttgc	gacaaaccac	aactagaatg	cagtaaaaaa	aatgctttat	2700
tttgtaaatt	tgtgtatgcta	ttgctttatt	tgttaaccatt	ataagctgca	ataaacaagt	2760
taacaacaac	aattgcattc	atTTTatgtt	tcaggttcag	ggggaggtgt	gggaggtttt	2820
ttaaagcaag	taaaacctct	acaaatgtgg	taaaatcgat	aaggatcaat	tcggcttcga	2880
ctagtactga	cggacacacc	gaagccccgg	cggcaaccct	cagcgatgc	cccggggctt	2940
cacgtttcc	caggtcagaa	gcggTTTcg	ggagtagtgc	cccaactggg	gtaacctttg	3000
agttctctca	gttggggcgc	tagggtcgcc	gacatgacac	aagggttgt	gaccgggtg	3060
gacacgtacg	cgggtgccta	cgaccgtcag	tcgcgcgagc	gcgactagta	caagccgaat	3120

Integrase.ST25

tgatccgtcg accgatgccc ttgagagcct tcaaccagt cagctccttc	3180
ggggcatgac tatcgctgcc gcacttatga ctgtcttctt tatcatgcaa	3240
aggtgccggc agcgctttc cgcttcctcg ctcactgact cgctgcgctc	3300
ctgcggcgag cggtatcagc tcactcaaag gcggtaatac gttatccac	3360
gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa	3420
ccgcgttgc tggcgaaaa ccataggctc cgccccctg acgagcatca	3480
cgctcaagtc agaggtggcg aaaccgaca ggactataaa gataccaggc	3540
ggaagctccc tcgtgcgctc tcctgttccg accctgcccgc ttaccggata	3600
tttctccctt cgggaagcgt ggcgcttct caatgctcac gctgttaggt	3660
gtgtaggtcg ttcgctccaa gctgggtgt gtgcacgaac ccccggttca	3720
tgcgccttat ccggtaacta tcgtcttgag tccaaacccgg taagacacga	3780
ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg	3840
ttcttgaagt ggtggctaa ctacggctac actagaagga cagtatttgg	3900
ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg	3960
accgctggta gcgggtggttt tttgtttgc aagcagcaga ttacgcgcag	4020
tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa	4080
cgttaaggga ttttggcat gagattatca aaaaggatct tcacctagat	4140
cattttaaatgaa gttttaatc aatctaaagt atatatgagt aaacttggc	4200
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctg	4260
tatttcgttc atccatagtt gcctgactcc ccgtcggtgt	4320
gataactacg atacggagg gcttaccatc tggccccagt	4380
gctgcaatga taccgcgaga cccacgctca cggctccag atttatcagc	4440
aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt	4500
tatccgcctc catccagtct attaattgtt gccggaaagc tagtaagt	4560
agttcgccag ttcatttgcgt cgtgggtgtca cgctcggt	4620
ttggatggc tccgggtccc aacgatcaag gcgagttaca tgatcccc	4680
tgttgcaaa aaaaacccgt gtcctccgat cttgtcaga agtaagttgg	4740
ccgcgttctcg agtcattctga gaatagtgtt tgccggcacc gagttgt	4800
cttgcgttgcgtt caatacggta taataccgcg ccacatagca	4860
gaactttaaa agtgcgtcatc attggaaaac gttttcgaa	4920
gcgaaaactc tcaaggatct taccgtgtt gagatccagt	4980
tcgatgtac ccactcggtc acccaactga tcttcagcat cttttacttt	5040
caccagcggtt tctgggtgag caaaaacagg aaggcaaaat gcccggaaaa	5100
aggaaataag ggcgacacgg aaatgttcaa tactcataact cttcctttt	5160
caatattattt gaagcattt tcagggttat tgtctcatga	
gcggatacat atttgaatgt atttggaaaa ataaacaaat aggggttccg	

Integrase.ST25

cgcacatttc cccgaaaagt gccacctgac	gcccctgta gcggcgcatt aagcgcgcg	5220
ggtgtggtgg ttacgcgcag cgtaccgct	acacttgcca gcccctagc gcccgtcct	5280
ttcgctttct tcccttcctt tctcgccacg	ttcgccggct ttccccgtca agctctaaat	5340
cgggggctcc cttaggggtt ccgatttagt	gctttacggc acctcgaccc caaaaaactt	5400
gattagggtg atggttcacg tagtgggcca	tcgcccgtat agacggtttt tcgccccttg	5460
acgttggagt ccacgttctt taatagtgga	ctcttggtcc aaactggaac aacactcaac	5520
cctatctcggtctt tgatttataa gggattttgc	cgatttcggc ctattggta	5580
aaaaatgagc tgatttaaca aaaatttaac	gcgaatttta acaaataatt aacgtttaca	5640
attccattt cggcatttcag gctgcgcaac	tgttggaaag ggcgatcggt gcgggcctct	5700
tcgctattac gccagccaa gctaccatga	taagtaagta atattaaggt acgggaggt	5760
cttggagcgg ccgcaataaa atatctttat	tttcattaca tctgtgtgtt gggtttttgt	5820
gtgaatcgat agtactaaca tacgctctcc	atcaaaacaa aacgaaacaa aacaaactag	5880
caaaataggc tgtccccagt gcaagtgcag	gtgccagaac attt	5924

<210> 4  
<211> 5101  
<212> DNA  
<213> Plasmid pCMV-pur-attB

<400> 4		
ctagagtcgg ggcggccggc cgcttcgagc	agacatgata agatacattt atgagtttgg	60
acaaaccaca actagaatgc agtaaaaaaaaa	atgctttattt tgtgaaattt gtgatgctat	120
tgctttattt gtaaccatta taagctgcaa	taaacaagtt aacaacaaca attgcattca	180
tttatgttt caggttcagg gggaggtgtt	ggaggttttt taaagcaagt aaaacctcta	240
caaatgttgtt aaaatcgata aggtcaatt	cggttcagg taccgtcgac gatgttaggtc	300
acggtctcga agccgcggtg cgggtgccag	ggcgtccct tgggctcccc gggcgctac	360
tccacctcac ccatctggtc catcatgatg	aacgggtcga ggtggcggtt gttgatcccgt	420
gcgaacgcgc ggcgcacccgg	gaagccctcg ccctcgaaac cgctggcgac	480
gtgagcacgg gacgtgcgac ggcgtcgacg	ggtgcggata cgcggggcag cgtcagcggg	540
ttctcgacgg tcacggcggtt	catgtcgaca gccgaattga tccgtcgacc	600
agagccttca acccagtca	gatgcccttg	660
cttatgactg tcttctttat	catgcaactc gtaggacagg tgccggcagc	720
ttcctcgctc actgactcgc	gatgtcggtt cggcgagcgg tatcagctca	780
ctcaaaggcg gtaatacggt	tatccacaga atcagggat aacgcaggaa	840
agcaaaaggc cagcaaaagg	agaacatgtg	900
ccaggaaccg taaaaaggcc	gcgttgctgg cgttttcca	960
taggctccgc cccccctgacg	agcatcacaa aaatcgacgc tcaagtca	1020
cccgacagga ctataaagat	ggtggcgaaa	1080
tgttccgacc ctgcccgtta	accaggcggtt tccccctgga agctccctcg	
ccggataacct gtccgcctt	tgcgctctcc	
ctcccttcgg	gaagcgtggc	

## Integrase.ST25

gctttctcaa	tgctcacgct	gtaggtatct	cagttcggtg	taggtcggtc	gctccaagct	1140
gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	1200
tctttagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	1260
gattagcaga	gcgaggtatg	taggcggtgc	tacagagttc	ttgaagtgg	ggcctaacta	1320
cggctacact	agaaggacag	tatgggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	1380
aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggttttt	1440
tgttgcaag	cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	cttgatctt	1500
ttctacgggg	tctgacgctc	agtggAACGA	aaactcacgt	taagggattt	tggcatgag	1560
attatcaaaa	aggatctca	cctagatcct	tttaaattaa	aatgaagtt	ttaaatcaat	1620
ctaaagtata	tatgagtaaa	cttggctctga	cagttaccaa	tgcttaatca	gtgaggcacc	1680
tatctcagcg	atctgtctat	ttcggtcatc	catagttgcc	tgactccccg	tcgtgttagat	1740
aactacgata	cgggagggct	taccatctgg	ccccagtgc	gcaatgatac	cgcgagaccc	1800
acgctcaccg	gctccagatt	tatcagcaat	aaaccagcca	gccggaaggg	ccgagcgcag	1860
aagtggcct	gcaactttat	ccgcctccat	ccagtcattt	aattgttgcc	gggaagctag	1920
agtaagtagt	tcgcccagtta	atagttgcg	caacgttgtt	gccattgcta	caggcatcgt	1980
ggtgtcacgc	tcgtcgtttgc	gtatggcttc	attcagctcc	ggttcccaac	gatcaaggcg	2040
agttacatga	tccccatgt	tgtcaaaaaa	agcggttagc	tccttcggc	ctccgatcgt	2100
tgtcagaagt	aagttggccg	cagtgttac	actcatggtt	atggcagcac	tgcataattc	2160
tcttactgtc	atgcccattccg	taagatgctt	ttctgtgact	ggtgagttact	caaccaagtc	2220
attctgagaa	tagtgtatgc	ggcgaccgag	ttgctttgc	ccggcgtcaa	tacggataaa	2280
taccgcgcca	catagcagaa	ctttaaaagt	gctcatcatt	ggaaaacgtt	cttcggggcg	2340
aaaactctca	aggatcttac	cgctgttgag	atccagttcg	atgtAACCCA	ctcgtgcacc	2400
caactgatct	ttagccatctt	ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	2460
gcaaaatgcc	gcaaaaaagg	gaataagggc	gacacggaaa	tgttgaatac	tcataacttt	2520
cctttttcaa	tattattgaa	gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	2580
tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	acattttccc	gaaaagtgcc	2640
acctgacgcg	ccctgttagcg	gcmcattaag	cgcggcggt	gtgggtggta	cgcgcagcg	2700
gaccgctaca	cttgcagcg	ccctagcgcc	cgctccttcc	gctttttcc	cttcctttct	2760
cgcacgttc	gccggcttc	cccgtaagc	tctaaatcgg	gggctccctt	tagggttccg	2820
atttagtgct	ttacggcacc	tcgacccaa	aaaacttgat	tagggtgatg	gttcacgtag	2880
tggccatcg	ccctgataga	cggttttcg	ccctttgacg	ttggagtcca	cgtttttaa	2940
tagtggactc	ttgttccaaa	ctggaacaac	actcaaccct	atctcggtct	attctttga	3000
tttataaggg	attttgccga	tttcggccta	ttggtaaaaa	aatgagctga	tttaacaaaa	3060
attaacgcg	aattttaca	aaatattaac	gtttacaatt	tcccattcgc	cattcaggct	3120

## Integrase.ST25

gcgcaactgt	tgccaagggc	gatcggtgcg	ggcctttcg	ctattacgcc	agcccaagct	3180	
accatgataa	gtaagtaata	ttaaggtacg	ggaggtactt	ggagcggccg	caataaaata	3240	
tctttat	ttt	cattacatct	gtgtgttgg	tttttgtgt	aatcgatagt	actaacatac	3300
gctctccatc	aaaacaaaac	gaaacaaaac	aaactagcaa	aataggctgt	ccccagtgca	3360	
agtgcaggtg	ccagaacatt	tctctatcga	tagtaccga	gctcttacgc	gtgctagccc	3420	
tcgagcagga	tctatacatt	gaatcaatat	tggcaattag	ccatattagt	cattggttat	3480	
atagcataaa	tcaatattgg	ctattggcca	ttgcatacgt	tgtatctata	tcataatatg	3540	
tacatttata	ttggctcatg	tccaatatga	ccgccatgtt	gacattgatt	attgactagt	3600	
tattaatagt	aatcaattac	gggtcatta	gttcatagcc	catatatgga	gttccgcgtt	3660	
acataactta	cgtaaatgg	ccgcctggc	tgaccgccc	acgaccccg	cccattgacg	3720	
tcaataatga	cgtatgttcc	catagtaacg	ccaataggga	ctttccattt	acgtcaatgg	3780	
gtggagtatt	tacggtaaac	tgcccactt	gcagtacatc	aagtgtatca	tatgccaagt	3840	
ccgcccccta	ttgacgtcaa	tgacggtaaa	tggccgcct	ggcattatgc	ccagtagatg	3900	
accttacggg	actttcctac	ttggcagtagc	atctacgtat	tagtcatcgc	tattaccatg	3960	
gtgatgcgg	tttggcagta	catcaatggg	cgtggatagc	ggtttgactc	acggggattt	4020	
ccaagtctcc	acccattga	cgtcaatggg	agtttggttt	ggcaccaaaa	tcaacgggac	4080	
tttccaaaat	gtcgtaacaa	ctccgc	ttgacgcaaa	tggccgtag	gcgtgtacgg	4140	
tggaggtct	atataagcag	agctcg	gtgaaccgtc	agatgcctg	gagacgccc	4200	
ccacgcttt	ttgacctcca	tagaagacac	cgggaccgat	ccagcctccc	ctcgaagctc	4260	
gactctagg	gctcgagatc	tgcgatctaa	gtaagcttgc	atgcctgcag	gtcggccgccc	4320	
acgaccgg	ccgcccaccat	cccgtacccc	acgcccctga	cccctcacaa	ggagacgacc	4380	
ttccatgacc	gagtacaagc	ccacgg	cctcgccacc	cgcgacgacg	tccccgggc	4440	
cgtacgcacc	ctcgccgccc	cgttgc	ctacccgc	acgcccaca	ccgtcgaccc	4500	
ggaccgcccac	atcgagcggg	tcaccgagct	gcaagaactc	ttcctcacgc	gcgtcggct	4560	
cgacatcg	aagggtgtgg	tgcggacga	cggccgcgc	gtggcggtct	ggaccacgccc	4620	
ggagagcgtc	gaagcgggg	cgggtttcgc	cgagatcggc	ccgcgcacgg	ccgagtttag	4680	
cgggtcccg	ctggccgcgc	agcaacagat	ggaaggcctc	ctggccgcgc	accggcccaa	4740	
ggagccgcgc	tgttcctgg	ccaccgtcg	cgtctgc	gaccaccagg	gcaagggtct	4800	
ggcagcgc	gtcgtgtcc	ccggagtgga	ggcggccgag	cgcgcgggg	tgccgcctt	4860	
cctggagacc	tccgc	gcaac	cttctacgag	cggctcg	tcaccgtcac	4920	
cgcgcacgtc	gaggtgc	aaggacgc	cacctgg	atgaccc	agcccggtgc	4980	
ctgacgccc	ccccacgacc	cgcacgc	gaccgaaagg	agcgcacgac	cccatggctc	5040	
cgaccgaagc	cgacc	ggccccc	acccgcacc	cgc	gcccaccgac	5100	
t						5101	

## Integrase.ST25

<210> 5  
 <211> 5043  
 <212> DNA  
 <213> Plasmid pCMV-pur-attP

<400> 5  
 ctagagtcgg ggcggccggc cgcttcgagc agacatgata agatacattg atgagtttgg 60  
 acaaaccaca actagaatgc agtaaaaaaaa atgctttatt tgtgaaattt gtgatgctat 120  
 tgctttattt gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca 180  
 ttttatgttt caggttcagg gggaggtgtg ggaggtttt taaagcaagt aaaacctcta 240  
 caaatgttgtt aaaaatcgata aggtcaatt cggcttcgac tagtactgac ggacacacccg 300  
 aagccccggc ggcaaccctc agcggatgcc ccggggcttc acgtttccc aggtcagaag 360  
 cggttttcgg gagtagtgcc ccaactgggg taacctttga gttctctcag ttggggcggt 420  
 aggtcgccg acatgacaca aggggttgg accgggggtgg acacgtacgc gggtgcttac 480  
 gaccgtcagt cgcgcgagcg cgactagtac aagccgaatt gatccgtcga ccgatgccct 540  
 tgagagcctt caacccagtc agtccttcc ggtggcgcg gggcatgact atcgtcgccg 600  
 cacttatgac tgtcttctt atcatgcaac tcgttaggaca ggtgccggca gcgctttcc 660  
 gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 720  
 cactcaaagg cgtaataacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 780  
 tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgttttc 840  
 cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 900  
 aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgtct 960  
 cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc ggaaagcgtg 1020  
 ggcgttttc aatgctcacg ctgttaggtat ctcagttcggt tgcgtccaaag 1080  
 ctgggctgtg tgacgaacc ccccggttcag cccgaccgct gcgccttac cggtaactat 1140  
 cgtcttgcgt ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac 1200  
 aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtgt gtggcctaac 1260  
 tacggctaca ctagaaggac agtattttgtt atctgcgttc tgctgaagcc agttaccttc 1320  
 ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggttt 1380  
 tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 1440  
 ttttctacgg ggtctgacgc tcagtggAAC gaaaactcac gttaaggat tttggtcgt 1500  
 agattatcaa aaaggatctt cacctagatc cttttaatt aaaaatgaag ttttaatca 1560  
 atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca 1620  
 cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgttag 1680  
 ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 1740  
 ccacgctcac cggtccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc 1800

Integrase.ST25

agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	ttaattgttg	ccggaaagct	1860	
agagtaagta	gttcgccagt	taatagtttg	cgcaacgttg	ttgccattgc	tacaggcatc	1920	
gtgggtcac	gctcgctg	ttgtatggct	tcattcagct	ccggttccca	acgatcaagg	1980	
cgagttacat	gatccccat	gttgcgaaa	aaagcggtta	gctccttcgg	tcctccgatc	2040	
gttgcagaa	gtaagttggc	cgcagtgta	tcactcatgg	ttatggcagc	actgcataat	2100	
tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	ctggtgagta	ctcaaccaag	2160	
tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	gcccgccgtc	aatacggat	2220	
aataccgcgc	cacatagcag	aactttaaa	gtgctcatca	ttggaaaacg	ttttcgggg	2280	
cgaaaactct	caaggatctt	accgctgttg	agatccagtt	cgatgttaacc	cactcgtgca	2340	
cccaactgat	cttcagcatc	ttttactttc	accagcg	ctgggtgagc	aaaaacagga	2400	
aggcaaaatg	ccgcaaaaaaa	gggataagg	gacacacgga	aatgttgaat	actcataactc	2460	
ttccttttc	aatattattt	aagcattt	cagggttatt	gtctcatgag	cggatacata	2520	
tttgaatgta	tttagaaaaaa	taaacaaata	ggggttccgc	gcacatttcc	ccgaaaagtg	2580	
ccacctgacg	cgcctgtag	cggcgcatta	agcgcggcgg	gtgtggtggt	tacgcgcagc	2640	
gtgaccgcta	cacttgccag	cgcctagcg	ccgc	tcgctttctt	cccttcctt	2700	
ctcgccacgt	tcgcccgtt	tccccgtcaa	gctctaaatc	ggggctccc	tttagggttc	2760	
cgatttagt	cttacggca	cctcgacccc	aaaaaactt	attagggtga	tggttcacgt	2820	
agtggccat	cgcctgata	gacggtttt	cgccttga	cgttggagtc	cacgttctt	2880	
aatagtggac	tcttgttcca	aactggaaca	acactcaacc	ctatctcggt	ctattcttt	2940	
gatttataag	ggattttgcc	gattcggcc	tattggtaa	aaaatgagct	gatttaacaa	3000	
aaatttaacg	cgaattttaa	caaaatatta	acgttacaa	tttccattc	gccattcagg	3060	
ctgcgcaact	gttggaaagg	gcgatcggtg	cggcctctt	cgctattacg	ccagcccaag	3120	
ctaccatgat	aagtaagta	tattaaggta	cgggaggtac	ttggagcggc	cgcaataaaa	3180	
tatcttatt	ttcattacat	ctgtgttt	gtttttgtg	tgaatcgata	gtactaacat	3240	
acgctctcca	tcaaaacaaa	acgaaacaaa	acaaactagc	aaaataggct	gtccccagtg	3300	
caagtgcagg	tgccagaaca	tttctctatc	gataggtacc	gagctttac	gcgtgctagc	3360	
cctcgagcag	gatctataca	ttgaatcaat	attggcaatt	agccatatta	gtcattggtt	3420	
atatacgata	aatcaatatt	ggctattggc	cattgcatac	gttgtatcta	tatcataata	3480	
tgtacattta	tattggctca	tgtccaatat	gaccgccatg	ttgacattga	ttattgacta	3540	
gttattaata	gtaatcaatt	acggggcat	tagttcatag	cccatatatg	gagttccgcg	3600	
ttacataact	tacggtaat	ggccgcctg	gctgaccgccc	caacgacccc	cgcatttga	3660	
cgtcaataat	gacgtatgtt	cccatagtaa	cgcacatagg	gactttccat	tgacgtcaat	3720	
gggtggagta	tttacggtaa	actgcccact	tggcagtaca	tcaagtgtat	catatgcca	3780	
gtccgc	cccccc	tattgacgtc	aatgacggta	aatggccgc	ctggcattat	gcccagtaca	3840

## Integrase(ST25)

tgaccttacg ggactttcct acttggcagt acatctacgt attagtcac	3900
tggatgatgcg gtttggcag tacatcaatg ggcgtggata gcggtttgc	3960
ttccaagtct ccacccatt gacgtcaatg ggagttgtt ttggcaccaa	4020
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca	4080
aatggcggt aggcgtgtac ggtgggaggt ctatataagc agagctcg	4140
tttgcacccatc catagaagac accgggaccg atccagcctc ccctcga	4200
agc tcgactctag gggctcgaga tctgcgatct aagtaagctt gcatgc	4260
ctgc agcccg tgccgccacc atcccctgac ccacgcccct gacc	4320
ccctcac aaggagacga cttccatga ccgagtacaa gcccacggtg	4380
cgccctcgca cccgcgacga cgtccccgg gccgtacgca ccctcgcc	4440
cgccgaccgc acatcgagcg ggtcaccgag ctgcaagaac tcttc	4500
cctc acatcg gcaagggtgtg ggtcgccgac gacggcgccg	4560
cggtggcggt ctggaccacg ccggagagcg tcgaagcgaa ggc	4620
gggtgttc gcagcaacag atggaaggcc tcctggcgcc gcacc	4680
ggccgcgc aaggagcccg cgtggttcct ggcaccgtc ggcgtctc	4740
cccgccgcgct ccccgagtg gaggcgccg agcgcgcgg ggtg	4800
ccctggaga cctccgcgcc ccgcaacctc cccttctacg agcgg	4860
ctcgacccgcg tcgaggtgcc cgaaggaccg cgacccgtt gcat	4920
gaccgcgcgc cggccacga cccgcagcgc ccgaccgaaa gg	4980
agcgcgcacg accccatggc tccgaccgaa gccgaccgg gc	5040
gggccccgc cgacccgcgca ccgcggccg aggcccaccg act	5043

&lt;210&gt; 6

&lt;211&gt; 5041

&lt;212&gt; DNA

&lt;213&gt; Plasmid pCMV-EGFP-attB

&lt;400&gt; 6

ctagagtcgg ggcggccggc cgcttcgagc agacatgata agata	60
tttttttttttgg acaaaccaca actagaatgc agtaaaaaaaaa	120
ttttttttttttttttaaaccatataaaccatgtttttttttttttttt	180
ttt	240
ttt	300
ttt	360
ttt	420
ttt	480
ttt	540
ttt	600

## Integrase.ST25

agagcttca acccagtcag ctccttcgg tggcgccgg gcatgactat cgtcgccgca	660
cttatgactg tcttcttat catgcaactc gtaggacagg tgccggcagc gctctccgc	720
ttcctcgctc actgactcgc tgcgctcggt cggtcggctg cggcgagcgg tatcagctca	780
ctcaaaggcg gtaatacggt tatccacaga atcagggat aacgcaggaa agaacatgtg	840
agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgaaaaa	900
taggctccgc cccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa	960
cccgacagga ctataaagat accaggcggt tccccctgga agctccctcg tgcgctctcc	1020
tgttccgacc ctgcccgtta ccggataacct gtccgcctt ctcccttcgg gaagcgtggc	1080
gctttctcaa tgctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct	1140
gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg	1200
tctttagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag	1260
gattagcaga gcgaggtatg taggcgggtgc tacagagtcc ttgaagtggg ggcctaacta	1320
cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg	1380
aaaaagagtt ggtagcttt gatccggcaa acaaaccacc gctggtagcg gtggttttt	1440
tgttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt	1500
ttctacgggg tctgacgctc agtggAACGA aaactcacgt taaggattt tggtcatgag	1560
attatcaaaa aggatctca cctagatcct tttaaattaa aaatgaagtt ttAAATCAAT	1620
ctaaagtata tatgagtaaa cttggctctga cagttaccaa tgcttaatca gtgaggcacc	1680
tatctcagcg atctgtctat ttctgttcatc catagttgcc tgactccccg tcgtgttagat	1740
aactacgata cgggagggt taccatctgg ccccaagtgc gcaatgatac cgcgagaccc	1800
acgctcaccc gctccagatt tatcagcaat aaaccagcca gccggaaagg ccgagcgcag	1860
aagtggctt gcaactttat ccgcctccat ccagtctatt aattgttgcc ggaaagctag	1920
agtaagtagt tcgcccgtta atagttgcg caacgttggt gccattgcta caggcatcg	1980
ggtgtcacgc tcgtcgttt gtatggcttc attcagctcc ggttcccaac gatcaaggcg	2040
atttacatga tccccatgt tgtcaaaaaa agcggttagc tccttcggc tcggatcg	2100
tgtcagaagt aagttggccg cagtgttatac actcatggtt atggcagcac tgcataattc	2160
tcttactgtc atgcccattcg taagatgctt ttctgtgact ggtgagtagt caaccaagtc	2220
attctgagaa tagtgtatgc ggcgaccgag ttgctttgc ccggcgtcaa tacggataaa	2280
taccgcgcca catagcagaa cttaaaagt gctcatcatt ggaaacgtt cttcggggcg	2340
aaaactctca aggatcttac cgctgttggatccatgatccatgatccatgatccatgatccat	2400
caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag	2460
gcaaaatgcc gcaaaaaagg gaataaggc gacacggaaa tggtaatac tcataactctt	2520
ccttttcaa tattattgaa gcatttatca gggattttgt ctcatgagcg gatacatatt	2580
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acattcccc gaaaagtgcc	2640

## Integrase.ST25

acctgacgct	ccctgttagcg	gcmcattaag	cgcggcggt	gtgggtgtt	cgcgcagcg	2700
gaccgctaca	cttgcgcagcg	ccctagcgcc	cgctccccc	gctttcttcc	cttcctttct	2760
cgcacgttc	gccggcttc	cccgtcaagc	tctaaatcgg	gggctccctt	tagggttccg	2820
atttagtgc	ttacggcacc	tcgacccaa	aaaacttgat	tagggtgatg	gttcacgtag	2880
tggccatcg	ccctgataga	cggttttcg	cccttgacg	ttggagtcca	cgttctttaa	2940
tagtggactc	ttgttccaaa	ctggaacaac	actcaaccct	atctcggtct	attcttttga	3000
tttataaggg	atttgccga	tttcggccta	ttggttaaaa	aatgagctga	tttaacaaaa	3060
atthaacgct	aattttaaca	aaatattaac	gtttacaatt	tcccattcgc	cattcaggct	3120
gcgcaactgt	tggaagggc	gatcggtgcg	ggcctttcg	ctattacgcc	agcccaagct	3180
accatgataa	gtaagtaata	ttaaggtacg	ggaggtactt	ggagcggccg	caataaaaata	3240
tctttat	cattacatct	gtgtgttgg	tttttgtgt	aatcgatagt	actaacatac	3300
gctctccatc	aaaacaaaaac	gaaacaaaaac	aaactagcaa	aataggctgt	ccccagtgca	3360
agtgcaggt	ccagaacatt	tctctatcga	tagtaccga	gctcttacgc	gtgctagccc	3420
tcgagcagga	tctatacatt	gaatcaatat	tggcaattag	ccatattagt	cattggttat	3480
atagcataaa	tcaatattgg	ctattggcca	ttgcatacgt	tgtatctata	tcataatatg	3540
tacattata	ttggctcatg	tccaatatga	ccgcatgtt	gacattgatt	attgactagt	3600
tattaatgt	aatcaattac	gggtcatta	gttcatagcc	catatatgga	gttccgcgtt	3660
acataactta	cgttaatgg	ccgcctggc	tgaccgccc	acgacccccc	cccattgacg	3720
tcaataatga	cgtatgtcc	catagtaacg	ccaataggga	cttccattg	acgtcaatgg	3780
gtggagtatt	tacggtaaac	tgcccactt	gcagtagatc	aagtgtatca	tatgccaagt	3840
ccgcccccta	ttgacgtcaa	tgacggtaaa	tggccgcct	ggcattatgc	ccagtagatg	3900
accttacgg	actttcctac	ttggcagtc	atctacgtat	tagtcatcgc	tattaccatg	3960
gtgatgcgtt	tttggcagta	catcaatggg	cgtggatagc	ggtttgactc	acggggattt	4020
ccaagtctcc	accccattga	cgtcaatggg	agttttttt	ggcaccaaaa	tcaacgggac	4080
tttccaaaat	gtcgtacaa	ctccgcccc	ttgacgcaa	tggcggtag	gcgtgtacgg	4140
tgggaggtct	atataagcag	agctcgaaa	gtgaaccgtc	agatcgctg	gagacgccat	4200
ccacgctgtt	ttgacctcca	tagaagacac	cgggaccgt	ccagcctccc	ctcgaagctc	4260
gactctagg	gctcgagatc	cccggttacc	ggtcgcccacc	atggtagca	agggcgagga	4320
gctgttacc	gggtgggtgc	ccatccttgt	cgagctggac	ggcgacgtaa	acggccacaa	4380
gttcagcgt	tccggcgagg	gcaaggcga	tgccacactac	ggcaagctga	ccctgaagtt	4440
catctgcacc	acggcaagc	tgcccgtgcc	ctggcccacc	ctcgtacca	ccctgaccta	4500
cggcgtgcag	tgcttcagcc	gctacccga	ccacatgaag	cagcacgact	tcttcaagtc	4560
cgcacatgccc	gaaggctacg	tccaggagcg	caccatcttc	ttcaaggacg	acggcaacta	4620
caagacccgc	gccgaggtga	agttcgaggg	cgacaccctg	gtgaaccgca	tcgagctgaa	4680

## Integrase.ST25

ggcatcgac ttcaaggagg acggcaacat cctggggcac aagctggagt acaactacaa	4740
cagccacaac gtctatatca tggccgacaa gcagaagaac ggcataagg tgaacttcaa	4800
gatccgcccc aacatcgagg acggcagcgt gcagctcgcc gaccactacc agcagaacac	4860
ccccatcgcc gacggcccg tgctgctgcc cgacaaccac tacctgagca cccagtccgc	4920
cctgagcaaa gaccccaacg agaagcgcga tcacatggc ctgctggagt tcgtgaccgc	4980
cgccgggatc actctcgca tggacgagct gtacaagtaa agcggccgct cgagcatgca	5040
t	5041

&lt;210&gt; 7

&lt;211&gt; 18116

&lt;212&gt; DNA

&lt;213&gt; Plasmid p12.01ys-LSPIPNNM-CMV-pur-attB

<400> 7	
ggctgcagg aattcgattt ccgccttctt tgatattcac tctgttgtat ttcatctttt	60
cttgcgcattt aaaggatata acagtctgta taacagtctg tgaggaaata cttggatttt	120
cttctgatca gtgtttttt aagtaatgtt gaatattgga taaggctgtg tgtcctttgt	180
cttggagac aaagcccaca gcaggtggtg gttgggtgg tggcagctca gtgacaggag	240
aggttttttt gcctgtttttt tttttttttt ttttttttaa gtaagggtttt cttttttttt	300
agtaaattttt ctactggact gtatgtttt acaggtcaga aacattttttt caaaagaaga	360
accttttggaa aactgtacag cccttttctt tcattccctt tttgctttctt gtgcattgc	420
cttggttctt gattgcatta tggaaaacgt tgatcgaaac ttgaggttttt tatttatagt	480
gtggcttgaa agcttggata gctgttggta cacgagatac cttatatagtt ttaggcccac	540
ttgatgctttt atttttttttt tttgaagtag tgagcgttctt ctggttttttt tcctttgaaa	600
ctgggtggc tttagattttt ctaatggat ttttacctg atgatctgt tgcataccac	660
aatgcttgta aatgttttcc tagttaacat gttgataact tcggatttac atgttgtata	720
tacttgcattt ctgtgtttctt agtaaaaaata tatggcattt atagaaatac gtaattcctg	780
atttcctttt ttttatctc tatgctctgt gtgtacaggt caaacagact tcactcctat	840
ttttatattt agaattttat atgcagtctg tcgttggttc ttgtgttggta aggatacagc	900
cttaaatttc ctagagcgat gctcagtaag gcgggttgc acatgggttc aaatgtaaaa	960
cggcacgtt tggctgctgc cttcccgaga tccaggacac taaactgctt ctgcactgag	1020
gtataaatcg cttcagatcc cagggaaatg cagatccacg tgcatttctt taaagaagaa	1080
tgaataactttt ctaaaatattt ttggcatagg aagcaagctg catggatttgg tttggactt	1140
aaattttttt ggtaacggag tgcataggtt taaaacacag ttgcagcatg ctaacgagtc	1200
acagcgtttta tgcagaagtg atgcctggat gcctgttgca gctgtttacg gcactgcctt	1260
gcagtgagca ttgcagatag ggggggggtg ctttgggtcg tggccaca cgctgcccaca	1320
cagccaccc tcgcacaca tctcacctgc tgggtacttt tcaaaccatc ttagcagtag	1380

Integrase\_ST25

tagatgagtt	actatgaaac	agagaagttc	ctcagttgga	tattctcatg	ggatgtcttt	1440
tttcccatgt	tggcaaagt	atgataaagc	atctctattt	gtaaattatg	cacttgttag	1500
ttcctgaatc	ctttctata	caccactt	tgcagcagg	gtaggctctg	gtgtggcctg	1560
tgtctgtgct	tcaatcttt	aaagcttctt	tggaaataca	ctgacttgat	tgaagtctct	1620
tgaagatagt	aaacagtact	taccttgat	cccaatgaaa	tcgagcattt	cagttgtaaa	1680
agaattccgc	ctattcatac	catgtaatgt	aattttacac	ccccagtgc	gacactttgg	1740
aatatattca	agtaatagac	tttggcctca	ccctcttgc	tactgtattt	tgtaatagaa	1800
aatatttaa	actgtgcata	tgattattac	attatgaaag	agacattctg	ctgatctca	1860
aatgtaagaa	aatgaggagt	gcgtgtgc	ttataaatac	aagtgattgc	aaatttagtgc	1920
aggtgtcctt	aaaaaaaaaa	aaaaaaagta	atataaaaag	gaccaggtgt	tttacaagtg	1980
aaatacattc	ctatttggta	aacagttaca	tttttatgaa	gattaccagc	gctgctgact	2040
ttctaaacat	aaggctgtat	tgtcttcctg	taccattgca	tttcctcatt	cccaatttgc	2100
acaaggatgt	ctggtaaac	tattcaagaa	atggcttga	aatacagcat	gggagcttgc	2160
ctgagttgga	atgcagagtt	gcactgcaaa	atgtcaggaa	atggatgtct	ctcagaatgc	2220
ccaactccaa	aggattttat	atgtgtat	agtaagcagt	ttcctgattc	cagcaggcca	2280
aagagtctgc	tgaatgtgt	gttgccggag	acctgtat	ctcaacaagg	taagatggta	2340
tcctagcaac	tgcggat	ttttttcag	cctttaggag	actgttaagc	aatttgcgt	2400
taggatcgt	ttcatcattt	tttagatgtt	tacttgaat	actgcataac	tttttagctt	2460
catgggttcc	ttttttcag	cctttaggag	actgttaagc	aatttgcgt	ccaaactttg	2520
tgttggtctt	aaactgcaat	agtagttac	cttgtattga	agaaataaag	accattttta	2580
tataaaaaaa	tactttgtc	tgtcttcatt	ttgacttgc	tgatatcctt	gcagtgc	2640
ttatgtcagt	tctgtcagat	attcagacat	caaaacttaa	cgtgagctca	gtggagttac	2700
tcataaaca	ttcatacag	agtgttaaggc	tagtgagaaa	tgcatacatt	tattgatact	2820
tttttaagt	caactttta	tcagat	tttcatttgc	gaaatataatt	gtttctaga	2880
ctgcatacgt	tctgaatctg	aaatgcagtc	tgattggcat	gaagaagcac	agcactttc	2940
atcttactta	aacttcattt	tggaaatgaag	gaagttaaagc	aagggcacag	gtccatgaaa	3000
tagagacagt	gcgcctcagga	gaaagtgaac	ctggatttct	ttggctagtg	ttctaaatct	3060
gtagtgagga	aagtaacacc	cgattccttgc	aaaggctcc	agcttaatg	cttccaaatt	3120
gaaggtggca	ggcaacttgg	ccactggta	tttactgcatt	tatgtctcag	tttcgcagct	3180
aacctggctt	ctccactatt	gagcatggac	tatgcctgg	cttcagaggc	caggtgaagg	3240
ttggatggg	tgaaaggagt	gctggctgt	ggctgggggg	actgtgggg	ctccaagctg	3300
agcttgggtt	ggcagcaca	ggaaaagtg	tggtaacta	tttttaagta	ctgtgttgca	3360
aacgtctcat	ctgcaaatac	gtagggtgt	tactctcgaa	gattaacagt	gtgggttcag	3420

Integrase.ST25

taatataatgg	atgaattcac	agtggaaagca	ttcaagggtta	gatcatctaa	cgacaccaga	3480
tcatcaagct	atgattggaa	gcggtatcag	aagagcgagg	aaggtaagca	gtcttcata	3540
gtttccctc	cacgtaaagc	agtctggaa	agtagcaccc	ctttagcaga	gacaaggaaa	3600
taattcagga	gcatgtgcta	ggagaacttt	cttgctgaat	tctacttgca	agagcttga	3660
tgcctggctt	ctgggtgcctt	ctgcagcacc	tgcaaggccc	agagcctgtg	gtgagctgga	3720
gggaaagatt	ctgctcaagt	ccaagcttca	gcaggtcatt	gtctttgctt	cttccccca	3780
cactgtgcag	cagagtggaa	ctgatgtcga	agcctcctgt	ccactacctg	ttgctgcagg	3840
cagactgctc	tcagaaaaag	agagctaact	ctatgccata	gtctgaaggt	aaaatgggtt	3900
ttaaaaaaaga	aaacacaaag	gcaaaaccgg	ctgccccatg	agaagaaagc	agtggtaaac	3960
atggtagaaa	aggtgcagaa	gcccccaggc	agtgtgacag	gccccctcctg	ccacctagag	4020
gcgggaacaa	gcttccctgc	ctagggctct	gcccgcgaag	tgcgtgtttc	tttgggtgggt	4080
tttggggc	gtttgggttt	gagatttaga	cacaaggaa	gcctgaaagg	aggtgttggg	4140
cactatttg	gtttgtaaag	cctgtacttc	aaatataat	tttgtgaggg	agtgtagcga	4200
attggccaat	ttaaaataaa	gttgcaagag	attgaaggct	gagtagttga	gagggttaaca	4260
cgttaatga	gatcttctga	aactactgct	tctaaacact	tgtttgagtg	gtgagacctt	4320
ggataggtga	gtgctttgt	tacatgtctg	atgcacttgc	ttgtcctttt	ccatccacat	4380
ccatgcattc	cacatccacg	catttgcac	ttatcccata	tctgtcatat	ctgacatacc	4440
tgtctttcg	tcacttggtc	agaagaaaca	gatgtgataa	tccccagccg	ccccaaagt	4500
gagaagatgg	cagttgcttc	tttccctttt	tcctgctaag	taaggatttt	ctcctggctt	4560
tgacaccta	cgaatagtc	ttcctgcctt	acattctggg	cattattca	aatatcttg	4620
gagtgcgctg	ctctcaagtt	tgtgtcttcc	tactcttaga	gtgaatgctc	ttagagtgaa	4680
agagaaggaa	gagaagatgt	tggccgcagt	tctctgatga	acacacctct	gaataatggc	4740
caaagggtgg	tgggtttctc	tgaggaacgg	gcagcgttt	cctctgaaag	caaggagctc	4800
tgcggagttg	cagttatttt	gcaactgatg	gtggaactgg	tgcttaaagc	agattcccta	4860
gttccctgc	tacttctttt	ccttcttggc	agtcagttt	tttctgacag	acaaacagcc	4920
acccccactg	caggcttaga	aagtatgtgg	ctctgcctgg	gtgtgttaca	gctctgccct	4980
ggtgaaaggg	gattaaaacg	ggcaccattc	atcccaaaca	ggatcctcat	tcatggatca	5040
agctgttaagg	aacttggct	ccaaccta	aacattaatt	ggagtacgaa	tgtaattaaa	5100
actgcattct	cgcattccta	agtcatttag	tctggactct	gcagcatgta	ggtcggcagc	5160
tcccactttc	tcaaagacca	ctgatggagg	agtagtaaaa	atggagaccg	attcagaaca	5220
accaacggag	tgttgccgaa	gaaactgatg	gaaataatgc	atgaattgtg	tgggtggacat	5280
ttttttaaa	tacataaaact	acttcaaatg	aggtcgagaa	aggtcagtgt	tttattagca	5340
gccataaaac	caggtgagcg	agtaccattt	ttctctacaa	gaaaaacgat	tctgagctct	5400
gcgttaagtat	aagttctcca	tagcggctga	agctcccccc	tggctgcctg	ccatctcagc	5460

Integrase(ST25)

tggagtgcag	tgccatttcc	ttggggtttc	tctcacagca	gtaatggac	aatactcac	5520
aaaaattctt	tctttcctg	tcatgtgg	tccctactgt	gccctcctgg	tttacgtt	5580
ccccctgact	gttccattca	gcgggttgg	aagagaaaaa	gaattggaa	ataaaacatg	5640
tctacgttat	cacccctcc	agcatttgg	tttttaatta	tgtcaataac	tggcttagat	5700
ttgaaatga	gaggggttg	ggtgtattac	cgaggaacaa	aggaaggctt	atataaactc	5760
aagtctttt	tttagagaac	tggcaagctg	tcaaaaacaa	aaaggcctt	ccaccaaatt	5820
aagtgaatag	ccgctatagc	cagcaggcc	agcacgaggg	atggtgact	gctggcacta	5880
tgccacggcc	tgcttgac	tctgagagca	actgctttgg	aatgacagc	acttggtgca	5940
atttccttt	tttcagaatg	cgtagagcgt	gtgcttggcg	acagttttc	tagttaggcc	6000
acttctttt	tccttctctc	ctcattctcc	taagcatgtc	tccatgctgg	taatcccagt	6060
caagtgaacg	ttcaaacaat	gaatccatca	ctgtaggatt	ctcggtgta	tcaaatctt	6120
gtgtgaggtc	tataaaat	ggaagcttat	ttattttcg	ttcttccata	tcagtcttct	6180
ctatgacaat	tcacatccac	cacagcaaat	taaaggtgaa	ggaggctggt	ggatgaaga	6240
gggtcttcta	gcttacgtt	cttccttgca	aggccacagg	aaaatgctga	gagctgtaga	6300
atacagcctg	ggtaagaag	ttcagtctcc	tgctggaca	gctaaccgca	tcttataacc	6360
ccttctgaga	ctcatcttag	gaccaaata	ggtctatctg	gggttttgt	tcctgctgtt	6420
cctcctggaa	ggctatctca	ctatccact	gctcccacgg	ttacaaacca	aagatacagc	6480
ctgaattttt	tctaggccac	attacataaa	ttgacctgg	taccaatatt	gttctctata	6540
tagttatttc	cttccccact	gtgttaacc	ccttaaggca	ttcagaacaa	ctagaatcat	6600
agaatggttt	ggattggaag	ggccttaaa	catcatccat	ttccaaccct	ctgccatggg	6660
ctgcttgcca	cccactggct	caggctgccc	agggccccat	ccagcctggc	ctttagcacc	6720
tccagggatg	gggcacccac	agttctctg	ggcagcctgt	gccaacacct	caccactctc	6780
tggtaaaga	attctctttt	aacatcta	ctaaatctct	tctcttttag	tttaaagcca	6840
ttcctctttt	tcccgttgct	atctgtccaa	gaaatgtgta	ttggtctccc	tcctgcttat	6900
aagcaggaag	tactggaagg	ctgcagtgag	gtctccccac	agccttctct	tctccaggct	6960
gaacaagccc	agtccttca	gcctgttttc	gttaggagatc	atcttagtgg	ccctcctctg	7020
gaccattcc	aacagttcca	cggtttctt	gtggagcccc	aggtctggat	gcagtacttc	7080
agatggggcc	ttacaaaggc	agagcagatg	gggacaatcg	cttacccttc	cctgctggct	7140
gccctgttt	tatgcagcc	caggtactg	ttggccttcc	aggctccag	acccttgct	7200
gatttgttc	aagctttca	tccaccagaa	cccacgcttc	ctggtaata	cttctgccc	7260
cacttctgt	agcttgtttc	aggagacttc	cattcttag	gacagactgt	gttacaccta	7320
cctgcccata	tcttgcatat	atacattca	gttcatgttt	cctgtaacag	gacagaat	7380
gtattcctct	aacaaaata	catgcagaat	tccttagtgcc	atctcagtag	ggttttcatg	7440
gcagtattag	cacatagtca	atttgctgca	agtacccccc	aagctgcggc	ctccataaa	7500

Integrase.ST25

tcctgtattt	gggatcagtt	accttttggg	gtaagctttt	gtatctgcag	agaccctggg	7560
ggttctgatg	tgcttcagct	ctgctctgtt	ctgactgcac	cattttctag	atcaccagg	7620
tgttcctgta	caacttcctt	gtcctccatc	cttcccagc	ttgtatcttt	gacaaataca	7680
ggcctatttt	tgtgttgct	tcagcagcca	ttaattctt	cagtgtcatc	ttgttctgtt	7740
gatgccactg	gaacaggatt	ttcagcagtc	ttgcaaagaa	catctagctg	aaaactttct	7800
gccattcaat	attcttacca	gttcttctt	tttgaggtga	gccataaatt	actagaactt	7860
cgtcaactgac	aagtttatgc	attttattac	ttcttattatg	tacttacttt	gacataaacac	7920
agacacgcac	atattttgct	gggatttcca	cagtgtctt	gtgtccttca	catggtttt	7980
ctgtcatact	tccgttataa	ccttggcaat	ctgcccagct	gcccatcaca	agaaaagaga	8040
ttcccttttt	attacttctc	ttcagccaat	aaacaaaatg	tgagaagccc	aaacaagaac	8100
tttgtgggca	ggctgccatc	aaggagaga	cagctgaagg	gttgtgtgc	tcaatagaat	8160
taagaaataa	taaagctgtg	tcagacagtt	ttgcctgatt	tatacaggca	cgccccaaac	8220
cagagaggct	gtctgccaag	gccaccttgc	agtccttgg	ttgtaagata	agtcataggt	8280
aactttctg	gtgaattgcg	tggagaatca	tgatggcagt	tcttgcgtt	tactatggta	8340
agatgctaaa	ataggagaca	gcaaagtaac	acttgctgct	gtaggtgctc	tgctatccag	8400
acagcgatgg	cactcgacac	ccaagatgag	ggatgctccc	agctgacgga	tgctggggca	8460
gtaacagtgg	gtcccatgct	gcctgctcat	tagcatcacc	tcagccctca	ccagccatc	8520
agaaggatca	tcccaagctg	aggaaagttg	ctcatttct	tcacatcatc	aaacctttgg	8580
cctgactgat	gcctcccgga	tgcttaatg	tggtcactga	catctttatt	tttctatgat	8640
ttcaagtcag	aacctccgga	tcaggaggga	acacatagtg	ggaatgtacc	ctcagctcca	8700
aggccagatc	ttccttcaat	gatcatgcat	gctacttagg	aaggtgtgt	tgtgtgaatg	8760
tagaattgcc	tttgttattt	tttcttctg	ctgtcaggaa	cattttgaat	accagagaaa	8820
aagaaaagtg	ctcttcttgg	catggagga	gttgcacac	ttgcaaaata	aaggatgcag	8880
tcccaaatgt	tcataatctc	agggtctgaa	ggaggatcag	aaactgtgt	tacaatttca	8940
ggcttctctg	aatgcagctt	ttgaaagctg	ttcctggccg	aggcagtaact	agtcagaacc	9000
ctcggaaaca	ggaacaaatg	tcttcaaggt	gcagcaggag	gaaacacctt	gcccatcatg	9060
aaagtgaata	accactgccg	ctgaaggaat	ccagctcctg	tttgagcagg	tgctgcacac	9120
tcccacactg	aaacaacagt	tcattttat	aggacttcca	ggaaggatct	tcttcttaag	9180
cttcttaatt	atggtacatc	tccagttggc	agatgactat	gactactgac	aggagaatga	9240
ggaactagct	ggaatattt	ctgtttgacc	accatggagt	cacccatttc	tttactggta	9300
tttggaaata	ataattctga	attgcaaagc	aggagttagc	gaagatctc	atttcttcca	9360
tgttgtgac	agcacagttc	tggctatgaa	agtctgctta	caaggaagag	gataaaaaatc	9420
ataggataa	taaatctaag	tttgaagaca	atgaggtttt	agctgcattt	gacatgaaga	9480
aattgagacc	tctactggat	agctatggta	tttacgtgtc	ttttgctta	gttacttatt	9540

Integrase.ST25

gaccccagct	gaggtcaagt	atgaactcag	gtctctcgaa	ctactggcat	ggattgatta	9600
catacaactg	taattttagc	agtgatttag	ggtttatgag	tactttgca	gtaaatcata	9660
gggttagtaa	tgttaatctc	aggaaaaaaa	aaaaaaagcc	aaccctgaca	gacatcccag	9720
ctcaggtgga	aatcaaggat	cacagctcag	tgcggtccc	gagaacacag	ggactcttct	9780
cttaggacct	ttatgtacag	ggcctcaaga	taactgatgt	tagtcagaag	actttccatt	9840
ctggccacag	ttcagctgag	gcaatcctgg	aattttctct	ccgctgcaca	gttccagtca	9900
tcccagttg	tacagttctg	gcacttttg	ggtcaggccg	tgatccaagg	agcagaagtt	9960
ccagctatgg	tcagggagtg	cctgaccgtc	ccaaactcact	gcactcaaac	aaaggcgaaa	10020
ccacaagagt	ggctttgtt	gaaattgcag	tgtggccag	aggggctgca	ccagtaactgg	10080
attgaccacg	aggcaacatt	aatcctcagc	aagtgcattt	tgcagccatt	aaattgaact	10140
aactgatact	acaatgcaat	cagtatcaac	aagtggttt	gcttggaaaga	tggagtctag	10200
gggctctaca	ggagtagcta	ctctctaatt	gagttgcatt	ttgaagcagg	acactgtgaa	10260
aagctggcct	cctaaagagg	ctgctaaaca	ttagggtcaa	ttttccagtg	cactttctga	10320
agtgtctgca	gttccccatg	caaagctgcc	caaacatagc	acttccaatt	gaatacaatt	10380
atatgcagggc	gtactgcttc	ttgccagcac	tgtccttctc	aaatgaactc	aacaaacaat	10440
ttcaaagtct	agtagaaagt	aacaagctt	aatgtcatt	aaaaagtata	tctgctttca	10500
gtagttcagc	ttatattatgc	ccactagaaa	catcttgtac	aagctgaaca	ctggggctcc	10560
agatttagtg	taaaacctac	tttatacaat	catagaatca	tagaatggcc	tgggttggaa	10620
gggaccccaa	ggatcatgaa	gatccaacac	ccccgccaca	ggcagggcca	ccaacctcca	10680
gatctggcac	tagaccaggc	agcccagggc	tccatccaac	ctggccatga	acacccctcag	10740
ggatggagca	tccacaacct	ctctggcag	cctgtgccag	cacccacca	ccctctctgt	10800
gaagaacttt	tccctgacat	ccaatctaag	ccttcctcc	ttgaggttag	atccactccc	10860
ccttgcata	tcactgtcta	ctctgtaaa	aagttgattc	tcctcctttt	tggaaggttt	10920
caatgaggcc	tccttgccagc	cttcttctct	tctgcaggat	gaacaagccc	agctccctca	10980
gcctgtctt	ataggagagg	tgctccagcc	ctctgatcat	ctttgtggcc	ctcctctgga	11040
cccgctccaa	gagctccaca	tcttcctgt	actgggggcc	ccaggcctga	atgcagttact	11100
ccagatgggg	cctcaaaaga	gcagagtaaa	gagggacaat	cacccctcc	accctgctgg	11160
ccagccctct	tctgatggag	ccctggatac	aactggcttt	ctgagctgca	acttctcctt	11220
atcagttcca	ctattaaac	aggaacaata	caacaggtgc	tgatggccag	tgcagagttt	11280
ttcacacttc	ttcatttcgg	tagatcttag	atgaggaacg	ttgaagttgt	gcttctgcgt	11340
gtgcttcttc	ctcctcaaat	actcctgcct	gatacctcac	cccacctgcc	actgaatggc	11400
tccatggccc	cctgcagcca	gggcctgat	gaacccggca	ctgcttcaga	tgctgtttaa	11460
tagcacagta	tgaccaagtt	gcacctatga	atacacaac	aatgtgttgc	atccttcagc	11520
acttgagaag	aagagccaaa	tttgcattgt	cagggaaatgg	tttagtaatt	ctgccaatta	11580

Integrase.ST25

aaacttgttt atctaccatg gctgtttta tggctgttag tagtggtaca ctgatgatga	11640
acaatggcta tgcagtaaaa tcaagactgt agatattgca acagactata aaattcctct	11700
gtggcttagc caatgtggta cttcccacat tgtataagaa atttggcaag ttttagagcaa	11760
tgtttgaagt gttggaaat ttctgtatac tcaagaggc gttttgaca actgtagaac	11820
agaggaatca aaaggggtg ggaggaagtt aaaagaagag gcaggtgcaa gagagcttgc	11880
agtcccgctg tgtgtacgac actggcaaca tgaggtctt gctaatttg gtgctttgct	11940
tcctgcccct ggctgcctta gggtgcgatc tgcctcagac ccacagcctg ggcagcagga	12000
ggaccctgat gctgctggct cagatgagga gaatcagcct gtttagctgc ctgaaggata	12060
ggcacgattt tggcttcct caagaggagt ttggcaacca gtttcagaag gctgagacca	12120
tccctgtgct gcacgagatg atccagcaga tctttaacct gtttagcacc aaggatagca	12180
gcgcgtctg gnatgagacc ctgctggata agtttacac cgagctgtac cagcagctga	12240
acgatctgga ggcttgcgtg atccagggcg tggcgtgac cgagaccct ctgatgaagg	12300
aggatagcat cctggctgtg aggaagtact ttcagaggat caccctgtac ctgaaggaga	12360
agaagtacag cccctgcgct tgggaagtcg tgaggcgtga gatcatgagg agcttagcc	12420
tgagcaccaa cctgcaagag agcttgaggt ctaaggagta aaaagtctag agtcggggcg	12480
gccggccgct tcgagcagac atgataagat acattgatga gtttggacaa accacaacta	12540
aatgcagtg aaaaaaatgc tttatttgc aaatttgc tgctattgct ttatttgc	12600
ccattataag ctgcaataaa caagttaca acaacaattt cattcatttt atgtttcagg	12660
ttcaggggaa ggtgtggag gtttttaaa gcaagtaaaa cctctacaaa tgtggtaaaa	12720
tcgataagga tccgtcgacc gatgccctt gagccttca acccagtcag ctccttcgg	12780
tggcgcggg gcatgactat cgtgcggca cttatgactg tcttctttat catgcaactc	12840
gtaggacagg tgccggcagc gctttccgc ttcctcgctc actgactcgc tgcgtcggt	12900
cgttcggctg cggcgagcgg tatcagctca ctcaaaaggcg gtaatacggt tatccacaga	12960
atcagggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg	13020
taaaaaggcc gcgttgctgg cgttttcca taggctccgc cccctgacg agcatcacaa	13080
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcg	13140
tccccctgga agctccctcg tgcgtctcc tggccgacc ctgcccctta ccggataacct	13200
gtccgcctt ctcccttcgg gaagcgtggc gctttctcaa tgctcacgct gtaggtatct	13260
cagttcggtg taggtcggtc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc	13320
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt	13380
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtgc	13440
tacagagttc ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat	13500
ctgcgtctg ctgaagccag ttaccttcgg aaaaagagtt ggtagcttt gatccggcaa	13560
acaaaccacc gctggtagcg gtggttttt tggcaag cagcagatta cgccgagaaa	13620

Integrase(ST25)

aaaaggatct	caagaagatc	cttgatctt	ttctacgggg	tctgacgctc	agtggAACGA	13680
aaactcacgt	taagggattt	tggcatgag	attatcaaaa	aggatcttca	cctagatcct	13740
tttaaattaa	aatgaagtt	ttaaatcaat	ctaaagtata	tatgagtaaa	cttggctga	13800
cagttaccaa	tgcttaatca	gtgaggcacc	tatctcagcg	atctgtctat	ttcgTTcatc	13860
catagttgcc	tgactccccg	tcgtgttagat	aactacgata	cgggaggggct	taccatctgg	13920
ccccagtgt	gcaatgatac	cgcgagaccc	acgctcaccg	gctccagatt	tatcagcaat	13980
aaaccagcca	gccggaaggg	ccgagcgcag	aagtggcct	gcaactttat	ccgcctccat	14040
ccagtctatt	aattgttgcc	gggaagctag	agtaagtagt	tcgCCAGTTA	atagttgCG	14100
caacgttgtt	gccattgcta	caggcatcgt	ggtgtcacgc	tcgtcgTTG	gtatggcttc	14160
attcagctcc	ggttcccaac	gatcaaggcg	atttacatga	tccccatgt	tgtcaaaaaa	14220
agcggttagc	tccttcggtc	ctccgatcgt	tgtcagaagt	aagttggccg	cagtgttATC	14280
actcatggtt	atggcagcac	tgcataattc	tcttactgtc	atgccatccg	taagatgctt	14340
ttctgtgact	ggtgagtagt	caaccaagtc	attctgagaa	tagtgtatgc	ggcgaccgag	14400
ttgctcttgc	ccggcgtcaa	tacggataa	taccgcGCCA	catagcagaa	ctttaaaagt	14460
gctcatcatt	ggaaaacgtt	cttcggggcg	aaaactctca	aggatcttac	cgctgttgag	14520
atccagttcg	atgtaaccca	ctcgtgcacc	caactgatct	tcagcatctt	ttactttcac	14580
cagcgTTCT	gggtgagcaa	aaacaggaag	gcaaaatGCC	gcaaaaaagg	gaataaggc	14640
gacacggaaa	tgttgaatac	tcatactctt	ccttttcaa	tattattgaa	gcatttatca	14700
gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	14760
ggttccgcgc	acatttcccc	gaaaagtGCC	acctgacgCG	ccctgttagcg	gcgcattaag	14820
cgcggcgggt	gtgggtggta	cgcgcagcgt	gaccgctaca	cttgcagcg	ccctagcgcc	14880
cgctcTTTC	gctttcttcc	cttccttct	cgccacgttc	gccggctttc	cccgtcaagc	14940
tctaaatcgg	gggctccctt	tagggtccg	atttagtgct	ttacggcacc	tcgaccccaa	15000
aaaacttgat	tagggtgatg	gttcacgtag	tggccatcg	ccctgataga	cggTTTTcg	15060
ccctttgacg	ttggagtcca	cgttctttaa	tagtggactc	ttgttccaaa	ctggaacaac	15120
actcaaccct	atctcggtct	attctttga	tttataaggg	attttGCCGA	tttcggccta	15180
ttggTTaaaa	aatgagctga	ttaacaaaa	atttaacgcg	aatttaaca	aaatattaac	15240
gtttacaatt	tcccattcgc	cattcaggct	gCGcaactgt	tgggaaggGC	gatcggtcg	15300
ggcctttcg	ctattacGCC	agcccaagct	accatgataa	gtaagtaata	ttaaggtacg	15360
ggaggtactt	ggagcggccg	ctctagaact	agtggatccc	ccggccgcaa	taaaatatct	15420
ttatTTcat	tacatctgtg	tgttggTTTT	ttgtgtgaat	cgatagtagt	aacatacgct	15480
ctccatcaaa	acaaaacgaa	acaaaacaaa	ctagcaaaaat	aggctgtccc	cagtgcAGT	15540
gcaggtGCCA	gaacatttct	ctatcgatag	gtaccgagct	cttacgcgtg	ctagccctcg	15600
agcaggatct	atacattgaa	tcaatattgg	caattagCCA	tattagtcat	tggTTatata	15660

Integrase.ST25

gcataaatca atattggcta ttggccattg catacggtt atctatatca taatatgtac 15720  
 atttatattg gctcatgtcc aatatgaccg ccatgttgac attgatttatt gactagttat 15780  
 taatagtaat caattacggg gtcattagtt catagcccat atatggagtt ccgcgttaca 15840  
 taacttacgg taaatggccc gcctggctga ccgcccacg acccccggcc attgacgtca 15900  
 ataatgacgt atgttcccat agtaacgcca atagggactt tccattgacg tcaatgggtg 15960  
 gagtatttac ggttaaactgc ccacttggca gtacatcaag tgtatcatat gccaagtccg 16020  
 ccccctattg acgtcaatga cggttaaatgg cccgcctggc attatgcccgttacatgacc 16080  
 ttacgggact ttcctacttg gcagttacatc tacgtatttag tcattcgctat taccatgggtg 16140  
 atgcggtttt ggcagttacat caatgggcgt ggatagcggt ttgactcacg gggattttcca 16200  
 agtctccacc ccattgacgt caatgggagt ttgtttggc accaaaatca acgggacttt 16260  
 cccaaatgtc gtaacaactc cgccccatttgc acgcaatgg gcggtaggcg tgtacgggtgg 16320  
 gaggtctata taagcagagc tcgttttagtgc aaccgtcaga tcgcctggag acgcccattca 16380  
 cgctgttttgc acctccatag aagacaccgg gaccgttcca gcctcccttc gaagctcgac 16440  
 tctaggggct cgagatctgc gatctaagta agcttgcatttgc cctgcaggcggc ggccgcccacg 16500  
 accgggtggccg ccaccatccc ctgaccacacg cccctgaccc ctcacaagga gacgaccccttc 16560  
 catgaccgag tacaagccca cggtgcgcct cgccacccgc gacgacgtcc cccggggccgt 16620  
 acgcaccctc gccgcccgt tcgcccacta ccccgccacg cgccacacccgc tcgaccggaa 16680  
 ccgcccacatc gagcgggtca ccgagctgca agaactcttc ctcacgcgcg tcgggctcga 16740  
 catcggtcaag gtgtgggtcg cggacgacgg cgccgcgggtg gcggtctggaa ccacgcccggaa 16800  
 gagcgtcgaa gcggggggcggttgcgatcgccgacgcgatggccg agttgagcggttccggc 16860  
 ttcccggtcg gccgcccgcgaaacagatggaaaggccttcgcgatggccgaccggccaaaggaa 16920  
 gcccgcgtgg ttccctggccca ccgtcggtcgcttcgcgatggccgaccggccaaaggaa 16980  
 cagcgcgcgtc gtgtcccgatggccgacgcgatggccgaccggccaaaggaa 17040  
 ggagacctcc gcgccccgcaccccttcctacgacgcgatggccgaccggccaaaggaa 17100  
 cgacgtcgag gtgtcccgatggccgacgcgatggccgaccggccaaaggaa 17160  
 acgccccccc cacgaccgcgacgcgccccgcacgcgatggccgaccggccaaaggaa 17220  
 ccgaagccga cccggggccgc cccgcgcaccgcgatggccgaccggccaaaggaa 17280  
 agagtcgggg cggccggccg cttcgagcagacatgataatgatggccgaccggccaaaggaa 17340  
 aaaccacaac tagaatgcag tgaaaaaaat gctttatttgc aatcattgtatggccgaccggccaaaggaa 17400  
 ctttatttgc aaccattata agctgcaata aacaagttaa caacaacaat tgcatttgc 17460  
 ttatgtttca ggttcagggg gaggtgtggg aggtttttaa aagcaagttaa aacctctaca 17520  
 aatgtggtaa aatcgataag gatcaattcg gcttcaggta ccgtcgacga tgttaggtcac 17580  
 ggtctcgaaag ccgcgggtcg ggtgccaggcgatggcccttg ggctccccgg ggcgtactc 17640  
 cacctcaccc atctggtcca tcatgatgaa cgggtcgagg tggcggttagt tgcatttgc 17700

## Integrase.ST25

gaacgcgcgg cgcaccggga agccctcgcc ctcgaaaccg ctggcgcg gg tggcacgg 17760  
 gagcacggga cgtgcgacgg cgtcgccggg tgcggatacg cggggcagcg tcagcgggtt 17820  
 ctcgacggtc acggcggca tgcgacagc cgaattgatc cgtcgaccga tgccctttag 17880  
 agccttcaac ccagtcagct cttccgggtg ggccggggc atgactatcg tcgcccact 17940  
 tatgactgtc ttctttatca tgcaactcgt aggacaggtg ccggcagcgc tcttccgctt 18000  
 cctcgctcac tgactcgctg cgctcggtc ttcggctgca gcgagcggta tcagctact 18060  
 caaaggcggt aatacggta tccacagaat cagggataa cgccaggaaag aacatg 18116

<210> 8  
 <211> 17402  
 <212> DNA  
 <213> Plasmid pOMIFN-Ins-CMV-pur-attB

<400> 8  
 ggccgcacc accgcgtggagc tccaattcgc cctatagtga gtcgtattac aattcactgg 60  
 ccgtcggtt acaacgtcgt gactggaaa accctggcgt taccctactt aatcgccctt 120  
 cagcacatcc cccttcgac agctggcgta atagcgaaga ggccgcacc gatgcgcctt 180  
 cccaaacgtt ggcgcgcgt aatggcgaat gggacgcgcctt ctgtagcggc gcattaaagcg 240  
 cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcctt ctagcgcctt 300  
 ctccttcgc tttcttcct tccttcgc ccacgttcgc cggcttcctt cgtcaagctc 360  
 taaatcgaaa gctccctta gggttccgat ttagtgctt acggcacctc gaccccaaaa 420  
 aacttgatta gggtgatggt tcacgttagtgc ggccatcgcc ctgatagacg gttttcgcc 480  
 cttgacgtt ggagtccacg ttcttaata gtggactctt gttccaaact ggaacaacac 540  
 tcaaccctat ctcggctat tctttgatt tataaggat tttgccatt tcggcctatt 600  
 ggttaaaaaa tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc 660  
 ttacaattta ggtggactt ttcggggaaa tgtgcgcgaa acccctattt gtttattttt 720  
 ctaaatacat tcaaataatgt atccgctcat gagacaataa ccctgataaaa tgcttcaata 780  
 atattgaaaa aggaagagta tgagtattca acattccgt gtcgcctta ttccctttt 840  
 tgcggcattt tgccttcctg ttttgctca cccagaaacg ctggtgaaag taaaagatgc 900  
 tgaagatcag ttgggtgcac gagtggtta catcgaactg gatctcaaca gcggtaaagat 960  
 ctttgcggcattt tttcgcccg aagaacgtt tccaatgttgc agcacttttta aagttctgt 1020  
 atgtggcgcg gtattatccc gtattgacgc cggcaagag caactcggtc gcccataca 1080  
 ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg 1140  
 catgacagta agagaattat gcagtgcgtc cataaccatg agtataaca ctgcggccaa 1200  
 cttacttctg acaacgatcg gaggaccgaa ggagcttacc gctttttgc acaacatgg 1260  
 ggatcatgta actcgcccttgc atcggtggaa accggagctg aatgaagcca taccacacg 1320  
 cgagcgtgac accacgatgc ctgttagcaat ggcaacaacg ttgcgcaaac tattaactgg 1380  
 cgaactactt actcttagtt cccggcaaca attaatagac tggatggagg cgatggaaagt 1440

## Integrase.ST25

tgcaaggacca	cttctgcgct	cggcccttcc	ggctggctgg	tttattgctg	ataaatctgg	1500
agccggtagag	cgtgggtctc	gcggtatcat	tgcagcactg	gggccagatg	gtaagccctc	1560
ccgtatcgta	gttatctaca	cgacggggag	tcaggcaact	atggatgaac	gaaatagaca	1620
gatcgctgag	ataggtgcct	cactgattaa	gcattggtaa	ctgtcagacc	aagtttactc	1680
atataactt	tagattgatt	taaaacttca	tttttaattt	aaaaggatct	aggtgaagat	1740
cctttttagat	aatctcatga	ccaaaatccc	ttaacgtgag	tttcgttcc	actgagcgtc	1800
agaccccgta	gaaaagatca	aaggatctc	ttgagatcct	tttttctgc	gcgtaatctg	1860
ctgcttgcaa	acaaaaaaaaac	caccgctacc	agcggtggtt	tgttgccgg	atcaagagct	1920
accaactctt	tttccgaagg	taactggctt	cagcagagcg	cagataccaa	atactgtcct	1980
tctagtgttag	ccgtagttag	gccaccactt	caagaactct	gtgcaccgc	ctacatacct	2040
cgctctgcta	atcctgttac	cagtggctgc	tgccagtgcc	gataagtcgt	gtcttaccgg	2100
gttggactca	agacgatagt	taccggataa	ggcgcagcgg	tcgggctgaa	cggggggttc	2160
gtgcacacag	cccagcttgg	agcgaacgac	ctacaccgaa	ctgagatacc	tacagcgtga	2220
gctatgagaa	agcgccacgc	ttcccgaagg	gagaaaggcg	gacaggtatc	cggtaagcgg	2280
cagggtcgga	acaggagagc	gcacgaggg	gcttccaggg	ggaaacgcct	ggtatctta	2340
tagtcctgtc	gggtttcgcc	acctctgact	tgagcgtcga	tttttgtat	gctcgtcagg	2400
ggggcggagc	ctatggaaaa	acgccagcaa	cgcggcctt	ttacggttcc	tggccttttgc	2460
ctggcctttt	gctcacatgt	tcttcctgc	gttatcccct	gattctgtgg	ataaccgtat	2520
taccgccttt	gagtgagctg	ataccgctcg	ccgcagccga	acgaccgagc	gcagcgtagtc	2580
agtgagcgtag	gaagcggaag	agcgccaaat	acgcaaaccg	cctctccccg	cgcgttggcc	2640
gattcattaa	tgcagctggc	acgacagggtt	tcccgactgg	aaagcgggca	gtgagcgc当地	2700
cgcaattaaat	gtgagttagc	tcactcatta	ggcacccag	gctttacact	ttatgcttcc	2760
ggctcgtatg	ttgtgtggaa	ttgtgagcgg	ataacaattt	cacacaggaa	acagctatga	2820
ccatgattac	gccaagctcg	aaattaaccc	tcactaaagg	gaacaaaagc	tgggtaccgg	2880
ccccccccc	gactagaggg	acagcccccc	cccaaagccc	ccagggatgt	aattacgtcc	2940
ctccccccgt	agggggcagc	agcgagccgc	ccggggctcc	gctccggtcc	ggcgctcccc	3000
ccgcatcccc	gagccggcag	cgtgcgggaa	cagccggc	acgggaaagg	tggcacggga	3060
tcgcttcct	ctgaacgctt	ctcgctgctc	ttttagcctg	cagacacctg	ggggatacg	3120
ggaaaaaagc	tttaggctga	aagagagatt	tagaatgaca	gaatcataga	acggcctggg	3180
ttgcaaagga	gcacagtgt	catccagatc	caacccctg	ctatgtcag	ggtcatcaac	3240
cagcagccca	ggctgcccag	agccacatcc	agcctggcct	tgaatgcctg	cagggatggg	3300
gcatccacag	cctccttggg	caacctgttc	agtgcgtcac	caccctctgg	ggaaaaaact	3360
gcctcctcat	atccaaccca	aacctcccct	gtctcagtgt	aaagccattc	ccccttgc当地	3420
tatcaagggg	gagtttgc当地	tgacattgtt	ggtctgggt	gacacatgtt	tgccaattca	3480

## Integrase.ST25

gtgcacacg gagaggcaga tcttgggat aaggaagtgc aggacagcat ggacgtggga	3540
catgcaggtg ttgagggctc tggacactc tccaaagtac acgcgttcaga acagccttaa	3600
ggataagaag ataggataga aggacaaaga gcaagttaaa acccagcatg gagaggagca	3660
caaaaaggcc acagacactg ctggcccttg tgtctgagcc tgcattttg atgggtctg	3720
gatgcaagca gaaggggtgg aagagcttgc ctggagagat acagctgggt cagtaggact	3780
gggacaggca gctggagaat tgccatgtag atgttcatac aatcgtaaaa tcatgaaggc	3840
tggaaaagcc ctccaagatc cccaaagacca accccaaccc acccaccgtg cccactggcc	3900
atgtccctca gtgccacatc cccacagttc ttcattcacct ccagggacgg tgacccccc	3960
acctccgtgg gcagctgtgc cactgcagca ccgcttttgg gagaaggtaa atcttgctaa	4020
atccagcccg accctccctt ggcacaacgt aaggccattt tctctcatcc aactccagga	4080
cggagtcagt gaggatgggg ctctagtcga ggtcgacggt atcgataagc ttgatttaggc	4140
agagcaatag gactctcaac ctcgtgagta tggcagcatg ttaactctgc actggagtcc	4200
agcgtggaa acaatctgcc ttgcacatga gtcttcgtgg gccaatattt cccaaacgggtt	4260
ttccttcagc ttgtcttgc tcctaagctc tcaaaacacc tttttggta ataaactcac	4320
ttggcaacgt ttatctgtct taccttagtg tcacgttca tccctattcc cctttctcct	4380
cctccgtgtg gtacacagtg gtgcacactg gttttctgt tgatgttctg ctctgacagc	4440
caatgtgggt aaagtttttc ctgccacgtg tctgtgttgc tttcacttca aaaaggccc	4500
tgggctcccc ttggagctt caggcatttc cttaatcatc acagtcacgc tggcaggatt	4560
agtccctcctt aaaccttata atgacctgaa cgtgtgtcc ctctttgttag tcagtgcagg	4620
gagacgtttg cctcaagatc agggtccatc tcacccacag ggccattccc aagatgaggt	4680
ggatggttta ctctcacaaa aagttttttt atgtttggct agaaaggaga actcaactgcc	4740
tacctgtgaa ttcccctagt cctgggttgc ctgccactgc tgcctgtgca gcctgtccc	4800
tggagggggc agcaactgct gtcacaaagg tgatcccacc ctgtctccac tgaaatgacc	4860
tcagtgccac gtgttgtata gggataaaag tacgggaggg ggatgcccgg ctcccttcag	4920
ggtgtcagag cagaagtgtc tgtgtataga gtgtgtctta atctattaat gtaacagaac	4980
aacttcagtc ctatgtttt gtgggctgga attgcccattt tgtagggac aggcttgcta	5040
aatcaactgca atcgccatag ttctgaaggt atttggaaa gaaaggatt tggggatttgc	5100
cctgtgattt gcttaattt gatggcaaat cacagaaag cagttctgtca caacagttgg	5160
ttgtttcagc caattcttgc agccaaagag ccgggtgccc agcgatataa tagttgtcac	5220
ttgtgtctgt atggatgaca gggaggtagg gtgacctgag gaccaccctc cagttctgc	5280
tagcgttaggt acagtcacca cctccagtc cacacgagtc ccatcgtggt ttaccaaaga	5340
aacacaatta ttggaccag ttggaaagt cacccgctga attgtgaggc tagattaata	5400
gagctgaaga gcaaatttgc ccaacttggaa gatacttagtt ggtatttagta tcagaggaac	5460
aggccatag cacctccatg ctattagatt ccggctggca tgtacttttca aagatgattt	5520

## Integrase.ST25

gtaactaaca atggcttatt	gtgcttgtct taagtctgtg	tcctaattgt aatgttcctt	5580
tggtttatat aaccttcttg	ccatttgc tc	ttcaggtgtt cttgcagaac	5640
ttaatctagt ttaactgtt	cttgattatt	cttagggata agatctgaat	5700
tggcttggc agactttagc	ttgggcttag	ctcccacatt agctttgct	5760
tgaagctatc aagatcctac	tcaatgacat	tagctgggtg caggtgtacc	5820
ctgtggaaca cattgtctga	tgataccgaa	ggcaaacgtg aactcaaaga	5880
taagaagaag tctgtcaat	tcagagaaaa	agccaaagtg gccattagac	5940
tgcagcattt gccagtaggt	ttcatataaa	actacaaaat ggaataaacc	6000
ggaaaagcct gatactagaa	tttaaatatt	cacccaggct caaggggtgt	6060
aatatcactc tataaaagta	ggcagccaa	ttattcacag acaaagctt	6120
tgctgcagt	ctgttttcg	gctgatccag ggttacttat	6180
gatttctcct tgtgtcatgt	tggtgaagga	gatatggcca gggggagatg	6240
agaggaaacg ttgcattttg	gtggcttggg	agaaaggtag aacgatatca	6300
gtcactaaga gatctgaagg	atggtttac	agaacagttg acttggctgg	6360
ggctgtaaat gnatggaagg	atggacagat	gggtggacag agattctgt	6420
atctcctgag ctcggtgctt	gacagactgc	agatccatcc cataaccttc	6480
gagcgcgggg agctttgta	ctgttcagtc	tgctgcttgt tgcttcctgg	6540
tgatattct tactcacaca	ggcaaaaac	ctgagcagct tcaaagtcaa	6600
tcataggcca ttcaagggtc	aagatgaggt	ttttggtttc ttgtttgtaa	6660
aagcactgaa gnatcagttg	cgagggcagg	gttttagcac tgttcagaga	6720
taactcctct catgaacaaa	aagagatgca	ggtgcagatt ctggcaagca	6780
gagaaagccc tgaatttctg	atatatgtgc	aatgttgggc acctaaccatt	6840
gcacagcagc tccagctcca	tgcagtactc	acagctggtg cagccctcgg	6900
tgagcagtgc tggactcac	gaggttccat	gtcttcaca ctgataatgg	6960
ggaatgggtg cccatccttg	gaggtcccc	aggccaggct ggctgcgtct	7020
cgatctggtg gtgagtagcc	agcccatggc	aggagttaga gcctgatggt	7080
ccttccaacc taagccatcc	tacgattcta	ggaatcatga cttgtgagtg	7140
aggcaatatt taaaagttat	aatgttttc	tccccttcct tgttgtcaa	7200
atcgccttat caatgcttt	ggagtctcca	gtcatttttc ttacacaaa	7260
agaatgaaga gaatcattta	atttcttgat	tgaatagtag gattcagaaa	7320
atgccgtctc tttgtatcga	gctgtaaggt	ttctcatcat ttatcagcgt	7380
agacttttc catctgatgt	ggaaaaaaaaa	atccttatca tctacagtct	7440
acatcgctca gactcttac	caaaaaagct	ataggtttta aaactacatc	7500
ttgccttgg	ttagctttc	ttccatatgc tgcgttgg	7560

## Integrase.ST25

aaactcttag	ctgctgagct	tgatgggtgc	ttaagaatga	agcactcact	gctgaaactg	7620
tttcatttc	acaggaatgt	tttagtggca	ttgttttat	aactacatat	tcctcagata	7680
aatgaaatcc	agaaataatt	atgcaaactc	actgcacccg	ttgcacaggt	ctttatctgc	7740
tagcaaagga	aataatttgg	ggatggcaaa	aacattcctt	cagacatcta	tattaaagg	7800
aatataatcc	tggtacccac	ccacttcatc	cctcattatg	ttcacactca	gagatactca	7860
ttctcttgtt	gttatcattt	gatagcgttt	tctttggttc	tttgcacgc	tctggctat	7920
ggctgcacgc	tctgcactga	ttagcaagta	gatgcgaggg	aagcagcagt	gagagggct	7980
gccctcagct	ggcacccagc	cgctcagcct	aggagggac	cttgccttc	caccagctga	8040
ggtgccagccc	tacaagctta	cacgtgctgc	gagcaggtga	gcaaagggag	tcttcatgg	8100
gtgtttcttg	ctgcccggaa	gcaaaacttt	actttcattc	attcccctt	aagaatgagg	8160
aatgtttgga	aacggactgc	tttacgttca	atttctctct	tcccttaag	gctcagccag	8220
gggccattgc	tgaggacggc	atcggggccc	cctggaccaa	atctgtggca	cagatggttt	8280
cacttacatc	agtggatgtg	gatctgcgc	ctgtaatgtg	tccttctgaa	ggaaggaacg	8340
tgccttccaa	gtgccagccc	cacagcccc	agcccctccc	tgtgctgctc	caattcatct	8400
cctcttcctc	tttctccctt	tgctgtttgt	gctcgggtag	aaatcatgaa	gatttagaag	8460
agaaaacaaa	ataactggag	tggaaaccca	ggtgatgcag	ttcattcagc	tgtcataggt	8520
ttgtcggtgc	tataggtctg	tatcagagat	gctarcacca	ctttgctgtc	ggtgcttaac	8580
tcgggtgaac	tctccttcac	tcgcatcatt	tgcggcctt	atttacatcc	ccagcatcca	8640
tcaccctctg	ggaaaatgg	cgcactggat	ctctaattgga	agactttccc	tcttcagag	8700
cctgtggat	gtgcagtgac	aagaaacgtg	gaggggctga	gcagcagcac	tgcccccagg	8760
gagcaggagc	ggatgccatc	ggtggcagca	tcccaaattga	tgtcagcgg	tgctgagcag	8820
gcagcggacg	aacggacaga	agcgatgcgt	acaccttctg	ttgacatggt	atttggcagc	8880
gatttaacac	tcgcttccta	gtcctgctat	tctccacagg	ctgcattcaa	atgaacgaag	8940
ggaaggagg	caaaaagatg	caaaatccga	gacaagcagc	agaaatattt	cttcgctacg	9000
gaagcgtgcg	caaacaacct	tctccaacag	caccagaaga	gcacagcgta	accttttca	9060
agaccagaaa	aggaaattca	caaagcctct	gtggatacca	gcgcgttcag	ctctcctgat	9120
agcagatttc	ttgtcagggtt	gcgaatgggg	tatggtgcca	ggaggtgcag	ggaccatatg	9180
atcatataca	gcacagcagt	cattgtgcat	gtattaatat	atattgagta	gcagtgttac	9240
tttgccaaag	caatagttca	gagatgagtc	ctgctgcata	cctctatctt	aaaactaact	9300
tataaatagt	aaaaccttct	cagttcagcc	acgtgctcct	ctctgtcagc	accaatggtg	9360
cttcgcctgc	acccagctgc	aaggaatcag	cccgtgatct	cattaacact	cagctctgca	9420
ggataaatta	gattgttcca	ctctctttt	ttgttaatta	cgacggaaca	attgttcagt	9480
gctgatggtc	ctaattgtca	gctacagaaa	acgtctccat	gcagttcctt	ctgcgccagc	9540
aaactgtcca	ggctatagca	ccgtgatgca	tgctacctct	cactccatcc	ttcttctt	9600

## Integrase.ST25

tcccaccagg gagagctgtg tggggactact ctcagccact ctgaacaata ccaaactgct	9660
acgcactgcc tccctcgaa agagaatccc cttgttgctt ttttatttac aggatccttc	9720
ttaaaaagca gaccatcatt cactgcaa ac ccagagcttc atgcctctcc ttccacaacc	9780
gaaaacagcc ggcttcattt gtctttta aatgctgttt tccaggtgaa ttttggccag	9840
cgtgttggct gagatccagg agcacgtgtc agcttctgc tctcattgct cctgttctgc	9900
attgcctctt tctggggtt ccaagagggg gggagacttt gcgcgggat gagataatgc	9960
ccctttctt agggtggctg ctggcagca gagtggctct gggtcactgt ggcaccaatg	10020
ggaggcacca gtgggggtgt gtttgtgca ggggggaagc attcacagaa tggggctgat	10080
cctgaagctt gcagtccaa gctttgtctg tgtacccagt gaaatccttc ctctgttaca	10140
taaagcccag ataggactca gaaatgttagt cattccagcc cccctttcc tcagatctgg	10200
agcagcactt gtttgcagcc agtcctcccc aaaatgcaca gacctcgccg agtggaggga	10260
gatgtaaaca gcgaaggta attacccct tgtcaaaaac actttgtggt ccatagatgt	10320
ttctgtcaat cttacaaaac agaaccgaga ggcagcgagc actgaagagc gtgttccat	10380
gctgagttaa tgagacttgg cagctcgctg tgcagagatg atccctgtgc ttcatggag	10440
gctgtaacct gtctcccat cgccttcaca ccgcagtgtc gtcctggaca ctcaccctc	10500
cataagctgt aggatgcagc tgcccaggaa tcaagagact tttcctaagg ctcttaggac	10560
tcatcttgc cgctcagtag cgtcagcaa ttactcatcc caactatact gaatgggtt	10620
ctgccagctc tgctgtttt tcaataagca tttcttcatt ttgcctctaa gtttctctca	10680
gcagcaccgc tctgggtgac ctgagtgcc acctggaacc cgagggcac agccaccacc	10740
tccctgttgc tgctgctcca gggactcatg tgctgctgaa tggggggaaag catgaagttc	10800
ctcaccaga cacctgggtt gcaatggctg cagcgtgctc ttcttggtat gcagattgtt	10860
tccagccatt acttgttagaa atgtgctgtg gaagcccttt gtatctctt ctgtggccct	10920
tcagcaaaag ctgtggaaa gctctgaggc tgctttttt ggtcgtggag gaattgtatg	10980
ttccttctt aacaaaaatt atccttagga gagagcactg tgcaagcatt gtgcacataa	11040
aacaattcag gttgaaaggg ctctctggag gtttccagcc tgactactgc tcgaagcaag	11100
gccaggttca aagatggctc aggatgctgt gtgccttcct gattatctgt gccaccaatg	11160
gaggagattc acagccactc tgctccctgt gccactcatg gagaggaata ttcccttata	11220
ttcagataga atgttatcct ttagctcagc cttccctata accccatgag ggagctgcag	11280
atccccatac tctcccttc tctgggtga aggccgtgtc ccccaaaaaa cttcccccacc	11340
ctgtgcccta agcagccgc tggcctctgc tggatgtgtg cctatatgtc aatgcctgtc	11400
cttgcagtcc agcctggac atttaattca tcaccaggat aatgtgaaac tgtgtcatct	11460
tcccctgcag ggtacaaagt tctgcacggg gtccttcgg ttcaggaaaa ctttcactgg	11520
tgctacactga atcaagctct atttaataag ttcataagca catggatgtg ttttcctaga	11580
gatacgaaaaatggatca gtgatttttta tttgctttgt tgcttacttc aaacagtgcc	11640

## Integrase.ST25

tttgggcagg	aggtgaggga	cgggtctgcc	gttggctctg	cagtgatttc	tccaggcgtg	11700
tggctcaggt	cagatagtgg	tcactctgtg	gccagaagaa	ggacaaagat	ggaaattgca	11760
gattgagtca	cgttaagcag	gcatcttgg	gtgattttag	gcagttcat	gaaagagcta	11820
cgaccactta	ttgttgtttt	cccctttac	aacagaagtt	ttcatcaaaa	taacgtggca	11880
aagcccagga	atgtttggga	aaagtgttagt	taaatgtttt	gtaattcatt	tgtcggagt	11940
ctaccagcta	agaaaaaaagt	cctaccttg	gtatggtagt	cctgcagaga	atacaacatc	12000
aatattagtt	tggaaaaaaa	caccaccacc	accagaaact	gtaatggaaa	atgtaaacca	12060
agaaattcct	tgggtaagag	agaaaggatg	tcgtatactg	gccaaagt	ccccagctgt	12120
cagcctgctg	accctctgca	gttcaggacc	atgaaacgtg	gcactgt	acgtgtcccc	12180
tgccttgct	tgcccacaga	tctctgcct	tgtgctgact	cctgcacaca	agagcattc	12240
cctgttagcca	aacagcgatt	agccataagc	tgcacctgac	tttgaggatt	aagagttgc	12300
aatTAAGTGG	attgcagcag	gagatcagt	gcagggttgc	agatgaaatc	cttttctagg	12360
ggtagctaag	ggctgagcaa	cctgtcctac	agcacaagcc	aaaccagcca	agggtttcc	12420
tgtgctgttc	acagaggcag	ggccagctgg	agctggagga	ggttgt	ggacccttct	12480
ccctgtgctg	agaatggagt	gatttctggg	tgctgttcc	gtggcttgca	ctgagcagct	12540
caagggagat	cggtgctcct	catgcagtgc	caaaactcgt	gtttgatgca	gaaagatgga	12600
tgtgcaccc	cctcctgcta	atgcagccgt	gagcttatga	aggcaatgag	ccctcagtgc	12660
agcaggagct	gtagtgcact	cctgttaggt	ctaggaaaaa	tctctgggtc	ccagggatgc	12720
attcataagg	gcaatataatc	ttgaggctgc	gccaaatctt	tctgaaatat	tcatgcgtgt	12780
tcccttaatt	tatagaaaca	aacacagcag	aataattatt	ccaatgcctc	ccctcgaagg	12840
aaacccat	ttccatgtag	aatgttaacc	tatatacaca	cagccatgct	gcattcattca	12900
gaacgtgcca	gtgctcatct	cccatggcaa	aatactacag	gtattctcac	tatgttggac	12960
ctgtgaaagg	aaccatggta	agaaacttcg	gttaaaggta	tggctgaaa	actactcata	13020
ccaaaacagc	agagctccag	acccctctt	aggaaagagc	cacttgaga	ggatgggt	13080
gaaggctgga	ggtgagagac	agagcctg	ccagtttcc	tgtctctatt	ttctgaaacg	13140
tttgcaggag	gaaaggacaa	ctgtactttc	aggcatagct	ggtgccctca	cgtaaataag	13200
ttccccgaac	ttctgtgtca	tttgttctta	agatgctttg	gcagaacact	ttgagtcaat	13260
tcgcttaact	gtgacttagt	ctgtaaataa	gtgctccctg	ctgataaggt	tcaagtgaca	13320
tttttagtgg	tatTTGACAG	catttacctt	gctttcaagt	cttctaccaa	gctcttctat	13380
acttaagcag	tgaaaccg	ccaaaccct	tcctttatc	aagctagtgc	taaataccat	13440
taacttcata	ggttagatac	ggtgctgcca	gcttcacctg	gcagtgg	ttgtcgttctg	13500
ctggtgacaa	agcctccctg	gcctgtgctt	ttaccttagag	gtgaatatcc	aagaatgcag	13560
aactgcattgg	aaagcagagc	tgcaggc	atggtgctga	gccttagctg	cttcctgctg	13620
ggagatgtgg	atgcagagac	gaatgaagga	cctgtccctt	actccctca	gcattctgt	13680

## Integrase.ST25

ctattttaggg ttctaccaga gtccttaaga ggttttttt tttttggc caaaagtctg	13740
tttgggttgg tttgaccact gagagcatgt gacacttgc tcaagctatt aaccaagtgt	13800
ccagccaaaa tcaattgcct gggagacgca gaccattacc tggaggtcag gacctaata	13860
aatattacca gcctcattgt gccgctgaca gattcagctg gctgctccgt gttccagtcc	13920
aacagttcgg acgccacgtt tgtatataatt tgcaggcagc ctcggggga ccatctcagg	13980
agcagagcac cgccagccgc ctgcagagcc gggcagtacc tcaccatggc tttgaccttt	14040
gccttactgg tggctctcct ggtgctgagc tgcaagagca gctgctctgt gggctgcgt	14100
ctgcctcaga cccacagcct gggcagcagg aggaccctga tgctgctggc tcagatgagg	14160
agaatcagcc tggtagctg cctgaaggat aggcacgatt ttggctttcc tcaagaggag	14220
tttggcaacc agtttcagaa ggctgagacc atccctgtgc tgcacgagat gatccagcag	14280
atcttaacc tggtagcac caaggatagc agcgctgctt gggatgagac cctgctggat	14340
aagtttaca ccgagctgta ccagcagctg aacgatctgg aggcttgcgt gatccaggc	14400
gtggcgtga ccgagacccc tctgatgaag gaggatagca tcctggctgt gaggaagtac	14460
ttcagagga tcaccctgta cctgaaggag aagaagtaca gcccctgcgc ttggaaagtc	14520
gtgaggcgt agatcatgag gagctttagc ctgagcacca acctgcaaga gagcttgagg	14580
tctaaggagt aaaaagtcta gagtcgggc ggccggccgc ttgcagcaga catgataaga	14640
tacattgatg agtttgaca aaccacaact agaatgcagt gaaaaaaatg ctttatttgc	14700
gaaatttgc atgctattgc ttatttgc accattataa gctgcaataa acaagttAAC	14760
aacaacaatt gcattcat tatgtttcag gttcaggggg aggtgtggg ggtttttaa	14820
agcaagtaaa acctctacaa atgtggtaaa atcgataccg tcgaccccgta ctagagcggc	14880
cactaacata cgctctccat caaaacaaaa cggaaacaaaa caaactagca aaataggctg	14940
tccccagtgc aagtgcaggt gccagaacat ttctctatcg ataggtaccg agctttacg	15000
cgtctagcc ctcgagcagg atctatacat tgaatcaata ttggcaatta gccatattag	15060
tcattggta tatagcataa atcaatattg gctattggcc attgcatacg ttgtatctat	15120
atcataatat gtacatttat attggctcat gtccaatatg accgccatgt tgacattgt	15180
tattgactag ttatataatg taatcaatta cggggtcatt agttcatagc ccatatatgg	15240
agttccgcgt tacataactt acggtaatg gcccgcctgg ctgaccgccc aacgacccccc	15300
gcccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt	15360
gacgtcaatg ggtggagttat ttacggtaaa ctgcccactt ggcagtcacat caagtgtatc	15420
atatgccaag tccgccccct attgacgtca atgacggtaa atggccgc tggcattatg	15480
cccagtacat gacccatcg gactttcta cttggcagta catctacgtt ttagtcatcg	15540
ctattaccat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact	15600
cacggggatt tccaagtctc cacccattt acgtcaatgg gagttgttt tggcaccaaa	15660
atcaacggga ctttccaaaa tgtcgtaaca actccgcctt attgacgcaa atggcggta	15720

## Integrase.ST25

ggcgtgtacg	gtgggaggc	tatataagca	gagctcg	ttt	agtgaaccgt	cagatgc	cct	15780					
ggagacgcca	tccacgctgt	tttgac	cctcc	atagaagaca	ccgggaccga	tccagc	cctcc	15840					
cctcgaagct	cgactctagg	ggctcg	agat	ctgcgatcta	agtaagctt	catgc	cctgc	15900					
ggtcggccgc	cacgaccggt	gccg	ccacca	tccc	ctgacc	cacg	cccctg	15960					
aggagacgac	cttccatgac	cgagtaca	ag	cccacgg	gcctcg	ccac	ccgac	16020					
gtcccc	ccgtac	gcac	cctcg	ccg	gcgtt	ccg	actac	cccc	16080				
accgtcgacc	cggacc	gcca	catcg	agcgg	gtcacc	gagc	tgcaaga	act	16140				
cgcgtcggc	tcgacat	cg	caagg	gtt	gg	acgg	ccgc	ggtgg	16200				
tggaccacgc	cggag	agcgt	cga	agcgggg	gcgg	tgtt	cg	ccgagat	cg	16260			
gccgagttga	gcgg	ttccc	g	ctgg	ccg	cag	caac	aga	tggagg	cct	16320		
caccggccca	aggag	ccc	gc	gtt	cct	gcc	accgt	gc	gtt	cc	16380		
ggcaagggtc	tggg	cag	gc	cgt	gtc	ccc	gg	agg	ccg	gc	16440		
gtgcccgc	cct	tg	ggag	ac	tc	cc	caac	ctt	ct	ac	16500		
ttcaccgtca	ccg	ccg	ac	gt	cg	cc	gagg	acc	gc	gtt	16560		
aagcccgg	cct	gac	gccc	gc	ccc	ac	gac	gc	ggaa	gag	16620		
ccccatgg	ccg	acc	gaa	gg	cc	ggcc	cc	acc	ccg	cc	16680		
ggcccacc	ctc	t	taga	gt	cc	ggcc	cc	gac	acat	ta	16740		
tgtgatgtt	gg	aca	aa	acc	ca	acta	ga	at	gt	aa	16800		
ttgtgatgt	att	gct	ttt	at	tg	taacc	tata	ag	ct	ttt	tgaa	16860	
caattgcatt	cattt	at	gt	t	ttc	aggtt	ca	gggg	agg	ttt	ttt	aa	16920
gtaaaac	ttc	tac	aaat	gt	gt	aaaat	cg	taa	agg	at	ttc	gg	16980
acgatgt	tc	ac	gg	t	tc	tc	ga	ag	cc	cg	tc	tc	17040
ccgggc	cg	tt	cc	tc	cc	cc	cc	cc	cc	cc	cc	cc	17100
tagttgat	cc	cgg	cg	ac	gc	gg	cc	gg	cc	cc	cc	cc	17160
gcgg	gtt	ca	gg	tc	gg	tc	gg	tc	gg	tc	gg	tc	17220
agcgtc	ag	gt	gg	tc	ac	gg	tc	gg	tc	gg	tc	gg	17280
ccgatgc	cc	t	g	at	gc	c	tt	cc	tt	cc	tt	cc	17340
atcgtcgcc	cac	tt	at	tg	ac	t	tt	tt	at	tg	ac	t	17400
gc													17402

<210> 9  
 <211> 5172  
 <212> DNA  
 <213> Plasmid pRSV-Int

<400> 9  
 ctgcattaaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgctttcc 60

Integrase(ST25)

gcttcctcgc	tcactgactc	gctgcgctcg	gtcggtcggc	tgcggcgagc	ggtatcagct	120
cactcaaagg	cggttaatacg	gttatccaca	aatcaggggg	ataacgcagg	aaagaacatg	180
tgagcaaaag	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgaaaa	240
cataggctcc	gccccccgtga	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	300
aacccgacag	gactataaag	ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	360
cctgttccga	ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	420
gcgctttctc	aatgctcacg	ctgttaggtat	ctcagttcgg	tgttaggtcgt	tcgctccaag	480
ctgggctgtg	tgcacgaacc	ccccgttcag	cccgaccgct	gcmccttac	cggtaactat	540
cgtcttgagt	ccaacccgtt	aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	600
aggattagca	gagcgaggta	tgttaggcgt	gctacagagt	tcttgaagtg	gtggcctaac	660
tacggctaca	ctagaaggac	agtatttggt	atctgcgctc	tgctgaagcc	agttaccttc	720
ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	cgggggtttt	780
tttggggca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	840
ttttctacgg	ggtctgacgc	tcagtggAAC	gaaaactcac	gttaaggat	tttggtcatg	900
agattatcaa	aaaggatctt	cacctagatc	cttttaaatt	aaaaatgaag	ttttaaatca	960
atctaaagta	tatatgagta	aacttggtct	gacagttacc	aatgcttaat	cagtgaggca	1020
cctatcttag	cgatctgtct	atttcggtca	tccatagttg	cctgactccc	cgtcgtag	1080
ataactacga	tacgggaggg	cttaccatct	ggccccagtg	ctgcaatgat	accgcgagac	1140
ccacgctcac	cggctccaga	tttatcagca	ataaaccagc	cagccggaag	ggccgagcgc	1200
agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	ttaattgtt	ccggaaagct	1260
agagtaagta	gttcgcccagt	taatagttt	cgcaacgtt	ttgccattgc	tacaggcatc	1320
gtgggtgtcac	gctcggtcg	ttgtatggct	tcatttagt	ccgggtccca	acgatcaagg	1380
cgagttacat	gatccccat	gttgtgcaaa	aaagcggta	gctccttcgg	tcctccgatc	1440
gttgtcagaa	gtaagttggc	cgcagtgtt	tcactcatgg	ttatggcagc	actgcataat	1500
tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	ctggtagt	ctcaaccaag	1560
tcattctgag	aatagtgtat	gcggcgaccg	agttgctt	gccccggcgtc	aatacggat	1620
aataccgcgc	cacatagcag	aactttaaa	gtgctcatca	ttggaaaacg	ttttcggggg	1680
cggaaaactct	caaggatctt	accgctgtt	agatccagtt	cgtgttaacc	cactcggtca	1740
cccaactgat	cttcagcatc	ttttactt	accagcgtt	ctgggtgagc	aaaaacagga	1800
aggcaaaaatg	ccgcaaaaaaa	gggataagg	gacacacgga	aatgttgaat	actcataact	1860
ttcccttttc	aatattattt	aagcattt	cagggttatt	gtctcatgag	cggtataata	1920
tttgaatgt	tttagaaaaaa	taaacaata	ggggttccgc	gcacatttcc	ccggaaaagt	1980
ccacctgacg	tcgacggatc	gggagatctc	ccgatccc	atggtcact	ctcagtacaa	2040
tctgctctga	tgccgcata	ttaagccagt	atctgcctcc	tgcttgcgt	ttggaggtcg	2100

Integrase\_ST25

ctgagtagtg	cgcgagcaaa	attnaagcta	caacaaggca	aggcttgacc	gacaattgca	2160
tgaagaatct	gcttagggtt	aggcgtttg	cgctgcttcg	cgatgtacgg	gccagatata	2220
cgcgtgctag	gggtcttagga	tcgattctag	gaattctcta	gccgcggtct	aggatcccgg	2280
gcgcgtatgg	tgcactctca	gtacaatctg	ctctgatgcc	gcatagttaa	gccagtatct	2340
gctccctgct	tgtgtgttgg	aggtcgctga	gtagtgcgcg	agcaaaattt	aagctacaac	2400
aaggcaaggc	ttgaccgaca	attgcatgaa	aatctgctt	agggttaggc	gtttgcgcgt	2460
gcttcgcgt	gtacgggcca	gatatacgcg	tatctgaggg	gactagggtg	tgttaggcg	2520
aaaagcgggg	cttcgggtgt	acgcggtag	gagtcccctc	aggatatagt	agttcgctt	2580
ttgcataggg	agggggaaat	gtagtcttat	gcaatacact	tgtagtcttgc	caacatggta	2640
acgatgagtt	agcaacatgc	cttacaagga	gagaaaaagc	accgtgcgt	ccgattgggt	2700
gaagtaaggt	ggtacgatcg	tgccttatta	ggaaggcaac	agacaggtct	gacatggatt	2760
ggacgaacca	ctgaattccg	cattgcagag	ataattgtat	ttaagtgcct	agtcgatac	2820
aataaacgcc	atttgcacat	tcaccacatt	ggtgtgcacc	tccaagcttgc	catgcctgca	2880
ggtaccggtc	cggaattccc	gggtcgacga	gctcaactgt	cgtagggtcg	ccgacatgac	2940
acaagggtt	gtgaccgggg	tggacacgta	cgcgggtgct	tacgaccgtc	agtcgcgcga	3000
gcgcgagaat	tcgagcgcag	caagcccagc	gacacagcgt	agcgccaacg	aagacaaggc	3060
ggccgacctt	cagcgcgaag	tcgagcgcga	cggggccgg	ttcaggttcg	tcgggcattt	3120
cagcgaagcg	ccgggcacgt	cggcggtcg	gacggcggag	cgcggaggt	tcgaacgcatt	3180
cctgaacgaa	tgccgcgccc	ggcggctcaa	catgatcatt	gtctatgacg	tgtcgcgctt	3240
ctcgccctg	aaggtcatgg	acgcgattcc	gattgtctcg	gaattgctcg	ccctgggcgt	3300
gacgatttt	tccactcagg	aaggcgcttt	ccggcaggga	aacgtcatgg	acctgattca	3360
cctgattatg	cggctcgacg	cgtcgacaa	agaatctcg	ctgaagtcgg	cgaagattct	3420
cgacacgaag	aactttcagc	gcgaattggg	cgggtacgtc	ggcgggaagg	cgccttacgg	3480
cttcgagctt	gttcggaga	cgaaggagat	cacgcgcaac	ggccgaatgg	tcaatgtcg	3540
catcaacaag	cttgcgcact	cgaccactcc	ccttaccgga	cccttcgagt	tcgagcccga	3600
cgtaatccgg	tggtggtggc	gtgagatcaa	gacgcacaaa	cacttccct	tcaagccggg	3660
cagtcaagcc	gccattcacc	cgggcagcat	cacggggctt	tgtaagcgca	tggacgctga	3720
cgcgtgccc	acccggggcg	agacgattgg	gaagaagacc	gcttcaagcg	cctggaccc	3780
ggcaaccgtt	atgcgaatcc	ttcgggaccc	gcgtattgcg	ggcttcgccc	ctgaggtgat	3840
ctacaagaag	aagccggacg	gcacgcccac	cacgaagatt	gagggttacc	gcattcagcg	3900
cgacccgatc	acgctccggc	cggtcgagct	tgattgcgg	ccgatcatcg	agcccgtga	3960
gtggtatgag	tttcaggcgt	gtttggacgg	cagggggcgc	ggcaaggggc	tttcccgggg	4020
gcaagccatt	ctgtccgcca	tggacaagct	gtactgcgag	tgtggcgccg	tcatgacttc	4080
gaagcgcggg	gaagaatcga	tcaaggactc	ttaccgctgc	cgtcgccgga	aggtggtcga	4140

## Integrase(ST25)

cccgtccgca cctgggcagc acgaaggcac gtgcaacgac agcatggcg cactcgacaa 4200  
 gttcggtcg gaacgcacatct tcaacaagat caggcacgac gaaggcgacg aagagacgtt 4260  
 ggcgcttctg tggaagccg cccgacgctt cggcaagctc actgaggcgc ctgagaagag 4320  
 cggcgaacgg gcgaacctt gtcggagcg cgccgacgac ctgaacgccc ttgaagagct 4380  
 gtacgaagac cgccggcag ggcgtacga cggaccgtt ggcaggaagc acttccggaa 4440  
 gcaacaggca gcgctgacgc tccggcagca agggcgaa gagcggctt ccgaacttga 4500  
 agccgccgaa gccccgaagc ttccccttga ccaatggttc cccgaagacg ccgacgctga 4560  
 cccgaccggc cctaagtcgt ggtggggcg cgcgtcagta gacgacaagc gcgtgtcgt 4620  
 cggctttc gtagacaaga tcgttgcac gaagtcact acggcagg ggcagggaaac 4680  
 gcccattcgag aagcgcgtt cgtacacgtg ggcgaagccg ccgaccgacg acgacgaaga 4740  
 cgacgcccac gacggcacgg aagacgtacg ggcgtacgac gacacccgg tccctcgagg 4800  
 ggccctattc tatagtgtca cctaaatgct agagctcgct gatcagcctc gactgtgcct 4860  
 tctagttgcc agccatctgt tggttgcctt tccccgtgc cttccttgac cctggaaggt 4920  
 gccactccca ctgtccttcc ctaataaaat gaggaaattt catgcattt tctgagtagg 4980  
 tgtcattcta ttctgggggg tgggtgggg caggacagca agggggagga ttggaaagac 5040  
 aatagcaggc atgctggggta tgccgtggc tctatggctt ctgaggcgg aagaaccagg 5100  
 tgcccagtca tagccgaata gcctctccac ccaagcggcc ggagaacctg cgtcaatcc 5160  
 actggggcg cg 5172

&lt;210&gt; 10

&lt;211&gt; 6233

&lt;212&gt; DNA

&lt;213&gt; Plasmid pCR-XL-TOPO-CMV-pur-attB

&lt;400&gt; 10

agcgcccaat acgcaaaccg cctctccccg cgccgtggcc gattcattaa tgcagctggc 60  
 acgacaggtt tcccgactgg aaagcggca gtgagcgca cgcaattaaat gtgagttac 120  
 tcactcatta ggcacccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa 180  
 ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac gccaagctat 240  
 ttaggtgacg cgttagaata ctcaagctat gcatcaagct tggtaccgag ctcggatcca 300  
 ctagtaacgg ccgcgcgtgt gctggattc gcccttggcc gcaataaaat atctttattt 360  
 tcattacatc tgtgtgttgg tttttgtgt gaatcgatag tactaacata cgctctccat 420  
 caaaacaaaa cggaaacaaaa caaactagca aaataggctg tccccagtgc aagtgcagg 480  
 gccaacat ttctctatcg ataggtaccg agctttacg cgtcttagcc ctcgagcagg 540  
 atctatacat tgaatcaata ttggcaatta gcatattag tcattggta tatagcataa 600  
 atcaatattt gctattggcc attgcatacg ttgtatctat atcataatat gtacatttt 660  
 attggctcat gtccaatatg accgccatgt tgacattgtat tattgactag ttatataat 720  
 taatcaatta cgggtcatt agttcatagc ccatatatgg agttccgcgt tacataactt 780

## Integrase.ST25

acggtaaatg	gcccgcctgg	ctgaccgccc	aacgacccccc	gcccattgac	gtcaataatg	840
acgtatgttc	ccatagtaac	gccaataggg	actttccatt	gacgtcaatg	ggtggagtat	900
ttacggtaaa	ctgcccactt	ggcagtgacat	caagtgtatc	atatgccaag	tccgccccct	960
attgacgtca	atgacggtaa	atggcccgcc	tggcattatg	cccagtgacat	gaccttacgg	1020
gactttccta	cttggcagta	catctacgta	ttagtcatcg	ctattaccat	ggtgatgcgg	1080
ttttggcagt	acatcaatgg	gcgtggatag	cggtttgact	cacggggatt	tccaagtctc	1140
caccccatgg	acgtcaatgg	gagtttgtt	tggcaccaaa	atcaacggga	ctttccaaaa	1200
tgtcgtaaca	actccgc(ccc)	attgacgcaa	atggcggta	ggcgtgtacg	gtgggaggtc	1260
tatataagca	gagctcgttt	agtgaaccgt	cagatcgct	ggagacgcca	tccacgctgt	1320
tttgacctcc	atagaagaca	ccgggaccga	tccagcctcc	cctcgaagct	cgactctagg	1380
ggctcgagat	ctgcgatcta	agtaagcttgc	catgcctgca	ggtcggccgc	cacgaccgg	1440
gccgccacca	tcccctgacc	cacgcccctg	accsctcaca	aggagacgac	cttccatgac	1500
cgagtacaag	cccacggtgc	gcctcgccac	ccgacgacgac	gtccccggg	ccgtacgcac	1560
cctcgccgcc	gcgttcgccc	actaccccgc	cacgcccac	accgtcgacc	cggaccgcca	1620
catcgagcgg	gtcaccgagc	tgcaagaact	cttcctcact	cgcgtcgggc	tcgacatcg	1680
caaggtgtgg	gtcgcggacg	acggcgccgc	ggtggcggtc	tggaccacgc	cgagagcgt	1740
cgaagcgggg	gcgggtttcg	ccgagatcg	cccgcgcatg	gccgagttga	gcggttcccg	1800
gctggcccg	cagcaacaga	tggaaggcct	cctggcgccg	caccggccca	aggagccgc	1860
gtggttcctg	gccaccgtcg	gcgtctcgcc	cgaccaccag	ggcaagggtc	tggcagcgc	1920
cgtcgtgctc	cccggagtgg	aggcgccga	gcgcgcccgg	gtgcccgcct	tcctggagac	1980
ctccgcgccc	cgcaacctcc	cttctacga	gcggctcgcc	ttcaccgtca	ccgcccacgt	2040
cgaggtgccc	gaaggaccgc	gcacctggtg	catgacccgc	aagcccgg	cctgacgccc	2100
gccccacgac	ccgcagcgcc	cgaccgaaag	gagcgcacga	ccccatggct	ccgaccgaag	2160
ccgaccgggg	cgccccgccc	gaccccgac	ccgcccccg	ggcccaccga	ctctagagtc	2220
ggggcggccg	ccgcgttcga	gcagacatga	taagatacat	tgtgatgttt	ggacaaacca	2280
caactagaat	gcagtgaaaa	aatgcttta	tttgtgaaat	ttgtgatgct	attgctttat	2340
ttgttaaccat	tataagctgc	aataaacaag	ttaacaacaa	caattgcatt	cattttatgt	2400
ttcaggttca	gggggaggtg	tgggaggttt	tttaaagcaa	gtaaaacctc	tacaaatgt	2460
gtaaaatcga	taaggatcaa	ttcggcttca	ggtaccgtcg	acgtatgtagg	tcacggtctc	2520
gaagccgcgg	tgcgggtgcc	agggcgtgcc	cttgggctcc	ccgggcgcgt	actccacctc	2580
accatctgg	tccatcatga	tgaacgggtc	gaggtggcgg	tagttgatcc	cggcgaacgc	2640
gcggcgcacc	ggaagccct	cgccctcgaa	accgctggc	gcgggtggta	cggtgagcac	2700
gggacgtgcg	acggcgtcgg	cgggtgcgga	tacgcggggc	agcgtcagcg	ggttctcgac	2760
ggtcacggcg	ggcatgtcga	cagccgaatt	gatccgtcga	ccgatgcct	tgagagcctt	2820

## Integrase.ST25

caacccagtc	agtccttcc	ggtggcgcg	gggcatgact	atcgctcgccg	cacttatgac	2880
tgtcttctt	atcatgcaac	tcgtaggaca	ggtgccggca	gcgccttcc	gcttcctcgc	2940
tcactgactc	gctgcgctcg	gtcggtcggc	tgcggcgagc	ggtatcagct	cactcaaagg	3000
cggttaatacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	aagggcgaat	3060
tctgcagata	tccatcacac	tggcggccgc	tcgagcatgc	atctagaggg	cccaattcgc	3120
cctatagtga	gtcgttattac	aattcactgg	ccgtcgttt	acaacgtcgt	gactggaaa	3180
accctggcgt	tacccaactt	aatgccttg	cagcacatcc	cccttcgccc	agctggcgt	3240
atagcgaaga	ggcccgacc	gatgcctt	cccaacagtt	gcgcagccta	tacgtacggc	3300
agtttaaggt	ttacacctat	aaaagagaga	gccgttatcg	tctgtttgtg	gatgtacaga	3360
gtgatattat	tgacacgccc	ggcgacgga	tggtgatccc	cctggccagt	gcacgtctgc	3420
tgtcagataa	agtctccgt	gaactttacc	cggtggtgca	tatcggggat	gaaagctggc	3480
gcatgatgac	caccgatatg	gccagtgtgc	cggctccgt	tatcggggaa	gaagtggctg	3540
atctcagcca	ccgcgaaaat	gacatcaaaa	acgccattaa	cctgatgttc	tgggaatat	3600
aaatgtcagg	catgagatta	tcaaaaagga	tcttcaccta	gatccttttc	acgtagaaag	3660
ccagtccgca	gaaacggtgc	tgaccccgga	tgaatgtcag	ctactggct	atctggacaa	3720
gggaaaacgc	aagcgcaaag	agaaagcagg	tagttgcag	tgggcttaca	tggcgatagc	3780
tagactggc	ggttttatgg	acagcaagcg	aaccggaatt	gccagctggg	gcgcctctg	3840
gtaaggttgg	gaagccctgc	aaagtaaact	ggatggctt	ctcgccgcca	aggatctgat	3900
ggcgcagggg	atcaagctt	gatcaagaga	caggtgagg	atcgtttgc	atgattgaac	3960
aagatggatt	gcacgcaggt	tctccggccg	tttgggtgga	gaggctattc	ggctatgact	4020
gggcacaaca	gacaatggc	tgctctgatg	ccgccgtgtt	ccggctgtca	gcgcaggggc	4080
gcccggttct	ttttgtcaag	accgacctgt	ccgggtccct	gaatgaactg	caagacgagg	4140
cagcgcggct	atcggtggct	gccacgacgg	gcgttccctg	cgcagctgtg	ctcgacgttg	4200
tcactgaagc	gggaaggac	tggctgctat	tggcgaagt	gccggggcag	gatctcctgt	4260
catctcacct	tgctcctgcc	gagaaagtat	ccatcatggc	tgatgcaatg	cgccggctgc	4320
atacgcttga	tccggctacc	tgcccattcg	accaccaagc	gaaacatcgc	atcgagcgag	4380
cacgtactcg	gatggaagcc	ggtcttgc	atcaggatga	tctggacgaa	gagcatcagg	4440
ggctcgcc	agccgaactg	ttcgccaggc	tcaaggcgag	catgcccgc	ggcgaggatc	4500
tcgtcgtac	ccatggcgat	gcctgcttgc	cgaatatcat	ggtggaaaat	ggccgtttt	4560
ctggattcat	cgactgtggc	cggtgggtg	tggcggaccg	ctatcaggac	atagcgttgg	4620
ctacccgtga	tattgctgaa	gagcttggcg	gcgaatggc	tgaccgcttc	ctcggtctt	4680
acggtatcgc	cgctcccgat	tcgcagcgca	tcgccttcta	tcgccttctt	gacgagttct	4740
tctgaattat	taacgcttac	aatttcctga	tgcggtattt	tctccttacg	catctgtgcg	4800
gtatccaca	ccgcatacag	gtggcacttt	tcggggaaat	gtgcgcggaa	cccctatttgc	4860

## Integrase.ST25

tttatttttc	taaatacatt	caaatatgta	tccgctcatg	agacaataac	cctgataaaat	4920
gcttcaataa	tagcacgtga	ggagggccac	catggccaag	ttgaccagtg	ccgttccgg	4980
gctcaccgcg	cgcgacgtcg	ccggagcggt	cgagttctgg	accgaccggc	tcgggttctc	5040
ccgggacttc	gtggaggacg	acttcgcccgg	tgtggtccgg	gacgacgtga	ccctgttcat	5100
cagcgcggtc	caggaccagg	tggtgccgga	caacaccctg	gcctgggtgt	gggtgcgcgg	5160
cctggacgag	ctgtacgccc	agtggtcgga	ggtcgtgtcc	acgaacttcc	gggacgcctc	5220
cgggcccggcc	atgaccgaga	tcggcgagca	gccgtggggg	cgggagttcg	ccctgcgcga	5280
cccgcccgcc	aactgcgtgc	acttcgtggc	cgaggaggcag	gactgacacg	tgctaaaact	5340
tcatttttaa	tttaaaagga	tctaggtgaa	gatcctttt	gataatctca	tgaccaaaat	5400
cccttaacgt	gagtttcgt	tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	5460
ttcttgagat	ccttttttc	tgcgcgtaat	ctgctgcttgc	caaacaaaaa	aaccaccgct	5520
accagcggtg	gtttgttgc	cggatcaaga	gctaccaact	cttttccga	aggtaactgg	5580
cttcagcaga	gcbcagatac	caaatactgt	ccttcttagtgc	tagccgtagt	taggccacca	5640
cttcaagaac	tctgttagcac	cgcctacata	cctcgctctg	ctaattctgt	taccagtggc	5700
tgctgccagt	ggcgataagt	cgtgtttac	cgggttggac	tcaagacgt	agttaccgga	5760
taaggcgcag	cggtcgggct	gaacgggggg	ttcgtgcaca	cagccagct	tggagcgaac	5820
gacctacacc	gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	cgcttcccga	5880
aggagaaaag	gcggacaggt	atccggtaag	cggcagggtc	ggaacaggag	agcgcacgag	5940
ggagcttcca	ggggaaacg	cctggtatct	ttatagtcct	gtcgggttcc	gccacctctg	6000
acttgagcgt	cgattttgt	gatgctcgtc	agggggcgg	agcctatgga	aaaacgcccag	6060
caacgcggcc	tttttacggt	tcctggcctt	ttgctggcct	tttgctcaca	tgttctttcc	6120
tgcgttatcc	cctgattctg	tggataaccg	tattaccgcc	tttgagttag	ctgataaccgc	6180
tcgcccgcagc	cgaacgaccg	agcgcagcga	gtcagtgagc	gaggaagcgg	aag	6233

<210> 11  
 <211> 234  
 <212> DNA  
 <213> artificial

<220>  
 <223> attP containing polynucleotide

<400> 11	gactagtact	gacggacaca	ccgaagcccc	ggcggcaacc	ctcagcggat	gccccggggc	60
	ttcacgtttt	cccaggtcag	aagcggtttt	cgggagtagt	gccccaaactg	ggtaacacctt	120
	ttagttctct	cagttggggg	cgtagggtcg	ccgacatgac	acaaggggtt	gtgaccgggg	180
	tggacacgta	cgcgggtgct	tacgaccgtc	agtgcgcgca	gcgcgactag	taca	234

<210> 12  
 <211> 26

## Integrase.ST25

<212> DNA  
<213> artificial

<220>  
<223> Primer attB-for

<400> 12  
taccgtcgac gatgtaggc acggtc

26